Front. Math. China 5:3 (2010), 379–515

SPEED OF STABILITY FOR BIRTH–DEATH PROCESSES

Mu-Fa Chen

(Beijing Normal University)

May 25, 2007—November 10, 2009

Abstract. This paper is a continuation of the study on the stability speed for Markov processes. It extends the previous study of the ergodic convergence speed to the non-ergodic one, in which the processes are even allowed to be explosive or to have general killings. At the beginning stage, this paper is concentrated on the birth-death processes. According to the classification of the boundaries, there are four cases plus one having general killings. In each case, some dual variational formulas for the convergence rate are presented, from which the criterion for the positivity of the rate and an approximating procedure of estimating the rate are deduced. As the first step of the approximation, the ratio of the resulting bounds is usually no more than 2. The criteria as well as basic estimates for more general types of stability are also presented. Even though the paper contributes mainly to the non-ergodic case, there are some improvements in the ergodic one. To illustrate the power of the results, a large number of examples are included.

Contents

1991 *Mathematics Subject Classification.* 60J27, 34L15.

Key words and phrases. Birth–death process, speed of stability, first eigenvalue, variational formula, criterion and basic estimates, approximating procedure, duality, killing.

Research supported in part by the Creative Research Group Fund of the National Natural Science Foundation of China (No. 10721091), by the "985" project from the Ministry of Education in China.

1. Introduction

Consider a birth–death process on the nonnegative integers $\mathbb{Z}_+$ with birth rates $b_n > 0\,(n \geqslant 0)$ and death rates $a_n > 0\,(n \geqslant 1)$. Define

$$\mu_0 = 1, \quad \mu_n = \frac{b_0 \cdots b_{n-1}}{a_1 \cdots a_n}, \qquad n \geqslant 1. \tag{1.1}$$

We say that the birth–death process is nonexplosive if the following Dobrushin's uniqueness criterion holds:

$$\sum_{k=0}^{\infty} \frac{1}{b_k \mu_k} \sum_{i=0}^{k} \mu_i \bigg[= \sum_{i=0}^{\infty} \mu_i \sum_{k=i}^{\infty} \frac{1}{b_k \mu_k} \bigg] = \infty \tag{1.2}$$

(cf. Dobrushin (1952), or Wang and Yang (1992, Corollary 5.2.1), or [10; Corollary 3.18]). This implies a useful condition that

$$\sum_{k=0}^{\infty} \bigg(\frac{1}{b_k \mu_k} + \mu_k \bigg) = \infty. \tag{1.3}$$

When $\sum_0^\infty \mu_k < \infty$, each of (1.2) and (1.3) is equivalent to the recurrent condition: $\sum_0^\infty (b_n \mu_n)^{-1} = \infty$. Otherwise, (1.3) cannot imply (1.2) since one can easily construct a counterexample so that $\sum_0^\infty \mu_k = \infty$ but

$$\sum_{i=0}^{\infty} \mu_i \sum_{k=i}^{\infty} \frac{1}{b_k \mu_k} < \infty.$$

Thus, under (1.3), the process may not be unique.

It is well known that for a birth–death process, the transition probabilities $(p_{ij}(t))$ satisfy

$$\lim_{t\to\infty} p_{ij}(t) =: \pi_j \geqslant 0 \tag{1.4}$$

for all $i, j \in \mathbb{Z}_+$. We are now interested in the exponential convergence rate

$$\alpha^* = \sup\Big\{ \alpha : \ |p_{ij}(t) - \pi_j| = O\big(\exp[-\alpha\, t]\big) \text{ as } t \to \infty \text{ for all } i, j \in E \Big\}. \tag{1.5}$$

In the ergodic case $\big($i.e., $\lim_{t\to\infty} p_{ij}(t) > 0$ for all i, $j\big)$, we have $Z := \sum_{j=0}^\infty \mu_j < \infty$ and then $\pi_j := \mu_j / Z > 0$ for all $j \geqslant 0$. In this case, the problem has been well studied, see, for instance, van Doorn (1981; 2002), Zeifman (1991), Kijima (1997), [2, 12], and the references therein. The problem becomes trivial in the zero-recurrent case for general irreducible Markov chains, since we have on the

one hand $\pi_j = 0$ for all j, and on the other hand, $\int_0^\infty p_{ii}(t)\mathrm{d}t = \infty$ for all i. Hence, the exponential decay can only happen in the transient case:

$$\sum_{n=0}^{\infty} \frac{1}{b_n \mu_n} < \infty. \tag{1.6}$$

Since the process is μ-symmetric: $\mu_i p_{ij}(t) = \mu_j p_{ji}(t)$ for all i, j and t, it is natural, as we did in the ergodic case, to use the L^2-theory. As usual, denote by $\|\cdot\|$ and $(\cdot,\cdot)$ the norm and the inner product on the real Hilbert space $L^2(\mu)$, respectively. Let

$$\mathscr{K} = \{f : f \text{ has finite support}\}. \tag{1.7}$$

Define

$$D(f) = \sum_{i\geqslant 0} \mu_i b_i (f_{i+1} - f_i)^2 = \sum_{i\geqslant 1} \mu_i a_i (f_i - f_{i-1})^2$$

with the minimal domain $\mathscr{D}^{\min}(D)$ consisting of the functions in the closure of $\mathscr{K}$ with respect to the norm $\|\cdot\|_D$: $\|f\|_D^2 = \|f\|^2 + D(f)$. Next, define

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f \in \mathscr{K}\} = \inf\{D(f) : \|f\| = 1, f \in \mathscr{D}^{\min}(D)\}.$$

From now on, we often write f_∞ or $f(\infty)$ as the limit of f at infinity provided it exists. In the definition of λ_0, it is natural to add the boundary condition $f_\infty = 0$ but this can be ignored since on the one hand, for each $f \in \mathscr{K}$, we have $f_\infty = 0$, and on the other hand $\mathscr{K}$ is a core of the Dirichlet form $\big(D, \mathscr{D}^{\min}(D)\big)$ (i.e., the form is regular) by [10; Proposition 6.59]. For a large part of the paper, we are dealing with this minimal Dirichlet form or the minimal process.

We now make a connection between α^* and λ_0. The proofs of the next three propositions are delayed for a moment.

Proposition 1.1. *For a general non-ergodic symmetric semigroup $\{P_t\}_{t\geqslant 0}$ with Dirichlet form $(D, \mathscr{D}(D))$ (not necessarily regular) on $L^2(\mu)$, the parameter λ_0,*

$$\lambda_0 = \inf\{D(f) : \|f\| = 1,\ f \in \mathscr{D}(D)\}, \tag{1.8}$$

is the largest ε such that

$$\|P_t f\| \leqslant \|f\|\, e^{-\varepsilon t}, \qquad t \geqslant 0, f \in L^2(\mu). \tag{1.9}$$

It was proved in [2; Theorem 5.3] that for birth–death processes, under (1.2), the exponentially ergodic convergence rate α^* coincides with the L^2-exponential one, denoted by λ_1:

$$\|P_t f - \pi(f)\| \leqslant \|f - \pi(f)\| e^{-\lambda_1 t} \qquad \text{for all } t \geqslant 0 \text{ and } f \in L^2(\mu),$$

where $\pi(f) = \int f \mathrm{d}\mu / \mu(E)$. For non-ergodic birth–death processes, we have similarly $\alpha^* = \lambda_0$, as mentioned at the end of [2]. Here is a generalization.

Proposition 1.2. *For a general non-ergodic μ-symmetric Markov chain with Dirichlet form $(D, \mathscr{D}(D))$, we have $\alpha^* = \lambda_0$ defined by (1.8).*

About (1.3), we have the following result.

Proposition 1.3. *Let $\mathscr{D}^{\max}(D) = \{f \in L^2(\mu) : D(f) < \infty\}$. Then the Dirichlet form $(D, \mathscr{D}^{\max}(D))$ is regular iff (1.3) holds. In other words, the Dirichlet form corresponding to the rates (a_i) and (b_i) is unique iff (1.3) holds.*

Proposition 1.2 reduces the study on α^* to the first (or principal) eigenvalue λ_0. This is the starting point of this paper. In the two cases we have discussed so far, the state 0 is a reflecting (Neumann) boundary, denoted by code "N". For λ_1, since the process starting from any point will certainly come back, the infinity may be regarded as a reflecting (Neumann) boundary. However, for λ_0, the situation is different. As we will prove in the next section, the corresponding eigenfunction decreases to zero at infinity. Hence, the infinity may be regarded as an absorbing (Dirichlet) boundary, denoted by code "D". Thus, for the temporary convenience, we rewrite $\lambda_1 = \lambda^{\rm NN}$ and $\lambda_0 = \lambda^{\rm ND}$. Replacing the Neumann boundary at 0 by the Dirichlet one (i.e., $b_0 = 0$), we obtain two more cases for which we have the decay rates (eigenvalues) $\lambda^{\rm DN}$ and $\lambda^{\rm DD}$, respectively. The main body of this paper is devoted to study these four cases. Now, the rate α^* coincides with, case by case, one of $\lambda^{\rm NN}$, $\lambda^{\rm ND}$, $\lambda^{\rm DN}$, and $\lambda^{\rm DD}$. Here are simple examples to show the difference in the different cases.

Examples 1.4.

(1) *Let $a_i = \delta i$, $b_i = \beta i + \gamma$, $\delta > \beta$. Then $\lambda^{\rm NN} = \delta - \beta$ if $\gamma > 0$ and so does $\lambda^{\rm DN}$ if $\gamma = 0$.*
(2) *Let $a_i = i$, $b_i = 2(i + \gamma)$. Then $\lambda^{\rm ND} = \gamma$ if $\gamma > 0$ and $\lambda^{\rm DD} = 1$ if $\gamma = 0$.*

The rate in the first example is the difference of the coefficients of leading terms, independent of γ. This is somehow natural. Surprisingly, the second one is determined by the constant term only except $\gamma = 0$ at which case there is a jump from $\lambda^{\rm ND}$ to $\lambda^{\rm DD}$. Thus, for the convergence rate, the role played by the parameters (a_i, b_i) is mazed and then one may wonder how far we can go (see Theorem 1.5 below for a preliminary answer).

The main body of the paper is devoted to the quantitive study of the convergence rate. For this, our key result (variational formulas) plays a full power. For those readers who are interested only in the qualitative criteria and basic estimates, here is a short statement.

Theorem 1.5 (Criterion and basic estimates). *Let (1.3) hold. Then in spite of $b_0 > 0$ or $b_0 = 0$, the exponential convergence rate α^* defined in (1.5) for the unique process is positive*

(1) *iff $\delta^{(4.4)} < \infty$ in the case of $\sum_i \mu_i < \infty$; and otherwise,*
(2) *iff $\delta^{(3.1)} < \infty$,*

where

$$\delta^{(4.4)} = \sup_{n \geqslant 1} \sum_{i=1}^{n} \frac{1}{\mu_i a_i} \sum_{j=n}^{\infty} \mu_j, \qquad \delta^{(3.1)} = \sup_{n \geqslant 0} \sum_{i=0}^{n} \mu_i \sum_{j=n}^{\infty} \frac{1}{\mu_j b_j}.$$

More precisely, we have the basic estimate $\delta^{-1}/4 \leqslant \alpha^* \leqslant \delta^{-1}$, where the constant δ is equal to $\kappa^{(6.13)}$ or $\kappa^{(7.5)}$ according to $b_0 > 0$ or $b_0 = 0$, respectively:

$$\big(\kappa^{(6.13)}\big)^{-1} = \inf_{m>n\geqslant 0}\bigg[\bigg(\sum_{i=0}^{n}\mu_i\bigg)^{-1} + \bigg(\sum_{i=m}^{\infty}\mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=n}^{m-1}\frac{1}{\mu_j b_j}\bigg)^{-1},$$

$$\big(\kappa^{(7.5)}\big)^{-1} = \inf_{m\geqslant n\geqslant 1}\bigg[\bigg(\sum_{i=1}^{n}\frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=m}^{\infty}\frac{1}{\mu_i b_i}\bigg)^{-1}\bigg]\bigg(\sum_{j=n}^{m}\mu_j\bigg)^{-1}.$$

Here, the superscript of $\kappa^{(7.5)}$, for instance, means that it is in the case studied in Section 7 and the constant is given in (7.5).

The proof of Theorem 1.5 and its extension are given in Section 7. The more general qualitative results are presented in Section 8 and in Summary 9.12 for the killing case.

To have an impression about the progress made in the paper, let us have a look at the new points made in the well-developed case, Section 6.

(1) The uniqueness condition (1.2) is replaced by using the maximal process.
(2) The more complete dual variational formulas are presented in Theorem 6.1.
(3) Even though the criterion Theorem 6.2 is known before, the upper bound in its improvement (Corollary 6.4) is newly added so that the ratio of the bounds is now no more than 2 as shown by a group of examples. Moreover, a new criterion (Corollary 6.6) which has been expected naturally (in view of Theorem 4.2) for a long time, is now presented.
(4) A more effective sequence for the upper estimate given in Theorem 6.3 is introduced to replace the original one. The monotonicity of the approximating sequences are proved here for the first time.

We now return to prove the propositions above.

Proof of Proposition 1.1. Replace by $\varepsilon_{\max}$ the largest exponential rate in (1.9). Then we have $\varepsilon_{\max} \geqslant 0$ because of the contractivity of the semigroup in every L^p-space ($p \geqslant 1$). We need to show that $\lambda_0 = \varepsilon_{\max}$. The proof of $\lambda_0 \geqslant \varepsilon_{\max}$ is easier since by an elementary property of the Dirichlet form and (1.9), we have for every f with $\|f\| = 1$,

$$D(f) = \lim_{t\downarrow 0}\uparrow \frac{1}{t}(f - P_t f,\, f) \geqslant \lim_{t\downarrow 0}\frac{1}{t}\big(1 - e^{-\varepsilon_{\max} t}\big) = \varepsilon_{\max}, \tag{1.10}$$

where $\lim\uparrow$ means an increasing limit. Hence, we have $\lambda_0 \geqslant \varepsilon_{\max}$. To prove $\varepsilon_{\max} \geqslant \lambda_0$, assume that $\lambda_0 > 0$. Otherwise, the assertion is trivial. Noticing that $D(f) = (-\Omega f,\, f)$ for the generator Ω of $\{P_t\}$ on $L^2(\mu)$ and for every $f \in \mathscr{D}(\Omega)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|P_t f\|^2 = 2(P_t f,\, \Omega P_t f) = -2D(P_t f). \tag{1.11}$$

Next, since $P_t f \in \mathscr{D}(D)$ for each $f \in L^2(\mu)$, by the definition of λ_0, we have

$$-2D(P_t f) \leqslant -2\lambda_0\|P_t f\|^2.$$

Thus, $\|P_t f\| \leqslant \|f\| e^{-\lambda_0 t}$ for all $t \geqslant 0$ and $f \in \mathscr{D}(\Omega)$, and then for all $f \in L^2(\mu)$ since the density of $\mathscr{D}(\Omega)$ in $L^2(\mu)$ and the contractivity of the semigroup $\{P_t\}_{t\geqslant 0}$. The assertion now follows since $\varepsilon_{\max}$ is the largest rate. $\square$

Proof of Proposition 1.2. The proof for $\alpha^* \geqslant \lambda_0$ is rather easy. Simply applying Proposition 1.1 to the indicator function $f = \mathbb{1}_{\{k\}}$, we obtain

$$p_{ik}(t) \leqslant \sqrt{\mu_k/\mu_i}\, e^{-\lambda_0 t}.$$

Note that this also provides a non-trivial estimate of the constant in (1.5).

To prove that $\lambda_0 \geqslant \alpha^*$, we may assume that $\alpha^* > 0$. One may follow the proof of [12; proof of part (4) of Theorem 8.13]. In the last part of the original proof, we have

$$\|P_t f\|^2 = (f, P_{2t} f) \leqslant \|f\|_\infty^2 e^{-2\alpha^* t} \sum_{i,j \in \mathrm{supp}(f)} \mu_i C_{ij}$$

for every bounded f with compact support. Here, we have used the assumption that $p_{ij}(t) \leqslant C_{ij} e^{-\alpha^* t}$. $\square$

Proof Proposition 1.3. Since the Q-matrix is conservative, by [10; Lemma 6.52 and Theorem 6.61], $(D, \mathscr{D}^{\max}(D))$ is a Dirichlet form and is indeed the maximal one. Note that in the conservative case, every Q-process (in particular, the semigroup generated by a Dirichlet form) satisfies the backward Kolmogorov's equation by [10; Theorem 1.15 (1)].

(a) Let (1.3) hold. Then the Dirichlet form should be regular. Otherwise, we have two different birth–death semigroups generated by $(D, \mathscr{D}^{\max}(D))$ and the minimal Dirichlet form $(D, \mathscr{D}^{\min}(D))$, respectively. They satisfy first the backward and then also the forward Kolmogorov's equations by [10; Theorem 6.16]. This is impossible since condition (1.3) is the uniqueness criterion for the process satisfying the Kolmogorov's equations simultaneously, due to Karlin and McGregor (1957a, Theorem 15) (cf. Hou et al. (2000, Theorem 6.4.6 (1); 1994, Theorem 12.7.1)). Note that criterion (1.3) is equivalent to the uniqueness for the process satisfying one of the Kolmogorov equations since every symmetric process as well as the minimal one satisfies both of the equations. This is the reason why (1.3) is weaker than (1.2).

(b) Next, let (1.3) fail. Then we have $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i b_i)^{-1} < \infty$. Moreover, (1.2) fails. Note that the birth–death Q-matrix has at most a single exit boundary, and there is precisely one if (1.2) fails. Besides, the non-trivial (maximal) exit solution z_λ is bounded from above by 1. In view of [10; Proposition 6.56], there are infinitely many Dirichlet forms. The minimal one is regular but not the maximal one $(D, \mathscr{D}^{\max}(D))$. $\square$

Actually, Proposition 1.3 is a particular case of a result we will study at the end of Section 9 (Theorem 9.22).

The remainder of the paper is organized as follows. In the next two sections, we study λ^{ND}. Sections 4, 6 and 7 are devoted to λ^{DN}, λ^{NN} and λ^{DD}, respectively. By exchanging N and D, we formally obtain a dual of λ^{ND} and λ^{DN} $\big($resp. λ^{NN} and $\lambda^{\mathrm{DD}}\big)$ which is studied in Section 5 (resp. 7). In each case, we present a group of dual variational formulas for the first (non-trivial) eigenvalue. By using the

formulas, we then deduce a criterion for the positivity of the eigenvalue and an approximating procedure for estimating the eigenvalue. The criteria and basic estimates in a quite general setup are given in Section 8. A closely related topic, having general killings, is studied in Section 9. In the study of this paper, the author has benefited a great deal from our previous work and from many authors' contribution. A part of the contributions is noted in the context. In the ergodic case under (1.2), a large number of references are given in [10, 12] and the author apologizes for omitting them here. At the end of the paper (Section 10), some remarks on the related results, some open problems or open topics, and so on are discussed. The analog of Theorem 1.5 for one-dimensional diffusions is also included.

Notation 1.6. *To be economical, we use the same notation λ_0, δ, κ, I and II and so on, from time to time in different sections with similar but different meaning. To distinguish them if necessary, we write $\lambda_0^{(\#)}$ for instance to denote the λ_0 defined by formula (#).*

2. Absorbing (Dirichlet) boundary at infinity: dual variational formulas

This section begins with the study on the property of eigenfunction of λ_0. It is fundamental in our analysis and has been studied several times before, see, for instance, [3; Lemma 4.2]; [4; proofs of Theorems 3.2 and 3.4]; Chen, Zhang and Zhao (2003, Section 2); Shao and Mao (2007, Proposition 3.1). The main body of this section is devoted to prove a group of variational formulas (Theorem 2.4 and Proposition 2.5). Their applications are given in the next section.

Fix an integer N: $1 \leqslant N \leqslant \infty$, and let $E = \{k \in \mathbb{Z}_+ : 0 \leqslant k < N+1\}$. Throughout the paper, the infinite case that $N = \infty$ is more essential but the finite case that $N < \infty$ is also included which may be meaningful in matrices theory. To avoid the confusion of these two cases in reading the paper, one may read the infinite case first and then go back to check the modification for the finite case. Besides, note that when $N < \infty$, neither (1.2) nor (1.3) is needed.

Let us start at a general situation. Consider the operator Ω corresponding to the birth–death Q-matrix with birth rates b_i, death rates a_i, and killing rates $c_i \geqslant 0\,(i \in E)$ as follows.

$$\Omega f(i)=b_i(f_{i+1}-f_i)+a_i(f_{i-1}-f_i)-c_i f_i,\quad i \in E,\ f_{N+1}=0 \text{ if } N<\infty. \tag{2.1}$$

In other words, when $N < \infty$, the state $N+1$ is an absorbing (Dirichlet) boundary. When $c_i \not\equiv 0$ for $1 \leqslant i < N$, unless otherwise stated, we assume that $a_0 = 0$ and $b_N = 0$ if $N < \infty$ (the other a_i and b_i are positive), otherwise, simply replace the original c_0 and c_N by $a_0 + c_0$ and $b_N + c_N$, respectively. Now, since $a_0 = 0$, f_{-1} is free in the last formula. The first eigenvalue λ_0 is now defined by

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f \in \mathscr{K}\}, \tag{2.2}$$

where

$$D(f) = \sum_{0\leqslant i<N} \mu_i b_i (f_{i+1} - f_i)^2 + \sum_{i\in E} \mu_i c_i f_i^2, \qquad f_{N+1} = 0 \text{ if } N < \infty. \tag{2.3}$$

We say that g is an "eigenfunction" of $\lambda \in \mathbb{R}$, if g satisfies the "eigenequation":

$$\Omega g = -\lambda g, \qquad g_{N+1} = 0 \text{ if } N < \infty. \tag{2.4}$$

Note that the "eigenvalue" and "eigenfunction" used in this paper are in a generalized sense rather than the standard ones since here we do not require $g \in L^2(\mu)$.

Proposition 2.1.

(1) *Every eigenfunction g of $\lambda \in \mathbb{R}$ satisfies*

$$\mu_k b_k (g_k - g_{k+1}) = \sum_{i=0}^{k} (\lambda - c_i)\mu_i g_i, \qquad k \in E,\ g_{N+1} = 0 \textit{ if } N < \infty. \tag{2.5}$$

(2) *If $\lambda_0 > 0$, then $c_i \not\equiv 0\,(0 \leqslant i \leqslant N)$ whenever $N < \infty$, and the non-zero eigenfunction g of λ_0 is either positive or negative on E.*

(3) *The non-zero eigenfunction g of $\lambda = 0$ is either positive and nondecreasing, or negative and nonincreasing on E. Furthermore, let $g > 0$ for instance. Then $g_{k+1} > g_k$ for all $k : i \leqslant k < N$ whenever $c_i > 0$.*

Proof. (a) Recall the eigenequation

$$\Omega g(i) = b_i(g_{i+1} - g_i) + a_i(g_{i-1} - g_i) - c_i g_i = -\lambda g_i, \qquad i \in E, \tag{2.6}$$

or more generally, the Poisson equation

$$b_i(g_i - g_{i+1}) - a_i(g_{i-1} - g_i) = h_i, \qquad i \in E,\ g_{N+1} = 0 \text{ if } N < \infty, \tag{2.7}$$

for a given function h. Multiplying both sides by μ_i, we get

$$\mu_i b_i(g_i - g_{i+1}) - \mu_{i-1} b_{i-1}(g_{i-1} - g_i) = \mu_i h_i, \qquad i \in E. \tag{2.8}$$

When $i = 0$, the second term on the left-hand side is set to be zero. Making a summation over i, we obtain

$$\mu_k b_k(g_k - g_{k+1}) = \sum_{i=0}^{k} \mu_i h_i, \qquad k \in E,\ g_{N+1} = 0 \text{ if } N < \infty. \tag{2.9}$$

With $h_i = (\lambda - c_i) g_i$, this gives us the first assertion of the proposition.

(b) To prove the second assertion, note that $\lambda_0 = 0$ if $c_i \equiv 0$ $(0 \leqslant i \leqslant N < \infty)$ in which case both 0 and N are reflecting and the process is ergodic. Now, since $\lambda_0 > 0$, one may assume that $g_0 \neq 0$, otherwise $g_i \equiv 0$ by induction. Next, replacing g by g/g_0 if necessary, we can assume that $g_0 = 1$. If g is not positive, then there would exist a $k_0 \in E$, $k_0 \geqslant 1$ such that $g_i > 0$ for $i < k_0$ and $g_{k_0} \leqslant 0$. We then modify g from k_0: set $\tilde{g}_i = g_i$ for $i < k_0$ and $\tilde{g}_i = 0$ for $i \geqslant k_0 + 1$. By choosing a suitable value $\varepsilon > 0$ at k_0, the new function $\tilde{g} \in \mathscr{K}$ gives us $D(\tilde{g})/\|\tilde{g}\|^2 < \lambda_0$, which is a contradiction to the definition of λ_0. Hence, g does not change its sign.

We are now going to specify ε. Note that

$$\begin{aligned}(-\Omega\tilde{g})(k_0-1) &= -b_{k_0-1}(\varepsilon - g_{k_0-1}) + a_{k_0-1}(g_{k_0-1} - g_{k_0-2}) + c_{k_0-1}g_{k_0-1}\\ &= (-\Omega g)(k_0-1) + b_{k_0-1}(g_{k_0} - \varepsilon)\\ &= \lambda_0 g_{k_0-1} + b_{k_0-1}(g_{k_0} - \varepsilon)\\ &< \lambda_0 g_{k_0-1}\end{aligned}$$

since $\varepsilon > 0 \geqslant g_{k_0}$. Note also that

$$(-\Omega\tilde{g})(k_0) = -b_{k_0}(0-\varepsilon) + a_{k_0}(\varepsilon - g_{k_0-1}) + c_{k_0}\varepsilon = \varepsilon(a_{k_0} + b_{k_0} + c_{k_0}) - a_{k_0}g_{k_0-1}.$$

Next, since $D(f) = (f, -\Omega f)$ for every $f \in \mathscr{K}$ and for each i, $\Omega f(i)$ depends on three points i and $i \pm 1$ only, we obtain

$$\begin{aligned}D(\tilde{g}) &= \sum_{0\leqslant i\leqslant k_0-2} \mu_i g_i(-\Omega g)(i) + \mu_{k_0-1}g_{k_0-1}(-\Omega\tilde{g})(k_0-1) + \mu_{k_0}\tilde{g}_{k_0}(-\Omega\tilde{g})(k_0)\\ &< \lambda_0 \sum_{i=0}^{k_0-1} \mu_i g_i^2 + \varepsilon\mu_{k_0}[\varepsilon(a_{k_0} + b_{k_0} + c_{k_0}) - a_{k_0}g_{k_0-1}].\end{aligned}$$

Because

$$\|\tilde{g}\|^2 = \sum_{i=0}^{k_0-1} \mu_i g_i^2 + \mu_{k_0}\varepsilon^2,$$

for $D(\tilde{g})/\|\tilde{g}\|^2 < \lambda_0$, it suffices that

$$\varepsilon[\varepsilon(a_{k_0} + b_{k_0} + c_{k_0}) - a_{k_0}g_{k_0-1}] < \lambda_0\varepsilon^2.$$

Equivalently,

$$\varepsilon(a_{k_0} + b_{k_0} + c_{k_0} - \lambda_0) < a_{k_0}g_{k_0-1}.$$

This clearly holds for sufficiently small $\varepsilon > 0$.

(c) If $\lambda = 0$, then (2.5) becomes

$$\mu_k b_k(g_{k+1} - g_k) = \sum_{i=0}^{k} c_i\mu_i g_i, \qquad k \in E,\ g_{N+1} = 0 \text{ if } N < \infty. \tag{2.10}$$

Clearly, if $g_0 = 0$, then $g_i \equiv 0$ by induction. Without loss of generality, assume that $g_0 = 1$. By (2.10) and induction, it follows that $g_{k+1} - g_k \geqslant 0$ for all $i \in E$. Actually, $g_{k+1} > g_k$ for all k: $i \leqslant k < N$ provided $c_i > 0$. $\square$

In view of (2.5), the eigenfunction g may not be monotone if $c_i \not\equiv 0$.

For the remainder of this section, we assume that $c_i = 0$ for $i < N$ but $c_N > 0$ if $N < \infty$. However, to simplify our notation, set $c_i \equiv 0$ but let $b_N > 0$ if $N < \infty$. In view of the definition of the state space E, the point $N+1$ is regarded as a Dirichlet boundary. From now on in the paper, when we talk about $\lambda_0^{(2.2)}$, it is defined by (2.2) *but in the present setting.*

Proposition 2.2.

(1) *Let g be a non-zero eigenfunction of $\lambda_0 > 0$. Then g is either positive or negative.*

(2) *Let g be a positive eigenfunction of $\lambda > 0$. Then g is strictly decreasing. Furthermore,*

$$\sum_{k=n}^{N} \frac{1}{\mu_k b_k} \sum_{i=0}^{k} \mu_i g_i = \frac{g_n - g_{N+1}}{\lambda}, \qquad n \in E. \tag{2.11}$$

In particular,

$$\sum_{n=0}^{N} \mu_n g_n \, \nu[n, N] = \sum_{n=0}^{N} \nu_n \sum_{k=0}^{n} \mu_k g_k = \frac{g_0 - g_{N+1}}{\lambda} < \infty, \tag{2.12}$$

where $\nu[\ell, m] = \sum_{k=\ell}^{m} \nu_k$, $\nu_k = (\mu_k b_k)^{-1}$. Moreover, if (1.2) holds, then $g_\infty := \lim_{N\to\infty} g_N = 0$.

(3) *Let $\lambda_0 = 0$. Then $N = \infty$ and the eigenfunction g must be a constant function.*

Proof. (a) The first assertion follows from Proposition 2.1 (2).

(b) Let $\lambda > 0$. Since $g > 0$, by (2.5) with $c_i \equiv 0$, it follows that g_i is strictly decreasing in i. By (2.5) again, we have

$$g_n - g_{N+1} = \sum_{k=n}^{N} (g_k - g_{k+1}) = \lambda \sum_{k=n}^{N} \frac{1}{\mu_k b_k} \sum_{i=0}^{k} \mu_i g_i = \lambda \sum_{i=0}^{N} \mu_i g_i \, \nu[i \vee n, N].$$

We obtain formula (2.11) and then (2.12). If $g_\infty > 0$, then by condition (1.2), the left-hand side of (2.11) is bounded below by

$$g_\infty \sum_{k=n}^{\infty} \frac{1}{\mu_k b_k} \sum_{i=0}^{k} \mu_i = \infty \tag{2.13}$$

which is a contradiction since the right-hand side of (2.11) is bounded from the above by $g_0/\lambda < \infty$. Therefore, we must have $g_\infty = 0$.

With some additional work, condition (1.2) for $g_\infty = 0$ will be removed (see Proposition 2.5 below).

(c) We now prove the last assertion of the proposition. When $N < \infty$, it is well known that $\lambda_0 > 0$. Now, let $\lambda_0 = 0$ and then $N = \infty$. By (2.6) with $c_i \equiv 0$, we have

$$g_{i+1} - g_i = \frac{a_i}{b_i} (g_i - g_{i-1}), \qquad i \geqslant 0.$$

From this and induction, it follows that $g_n = g_0$ for all $n \geqslant 1$ since $a_0 = 0$. $\square$

We remark that for finite state space with absorbing at $N+1 < \infty$, Proposition 2.2 was actually proved in [4; proof d) of Theorem 3.4] with a change of the order of the state space. Next, when $N = \infty$ and $\lambda_0 > 0$, in contrast with the ergodic case where $g \in L^1(\mu)$ (cf. [12; Proposition 3.5]), here one may have $g \notin L^2(\mu)$ and then $g \notin L^1(\mu)$. However, $g \in L^1(\nu)$ since g_n is strictly decreasing and $\sum_n \nu_n < \infty$, which is a consequence of Theorem 3.1 below.

Corollary 2.3. *Let* $\lambda_0 > 0$. *Then* $\lim_{i\to\infty} P_t f(i) = 0$ *for all* $t \geqslant 0$ *and* $f \in \mathscr{K}$.

Proof. It suffices to show that $\lim_{i\to\infty} \sum_{k=1}^n p_{ik}(t) = 0$. We now prove a stronger conclusion: $\lim_{i\to\infty} P_t g(i) = 0$ for all $t \geqslant 0$, where $g > 0$ with $g_0 = 1$ is the eigenfunction of λ_0. Since g is bounded, by using the well-known fact that

$$e^{-\lambda_0 t} g_i = P_t g(i), \qquad t \geqslant 0,$$

the conclusion now follows from Propositions 2.2 and 2.5 (2) below. □

For a specialist who does not want to know many details, at the first reading, one may have a glance at the remainder of this section and the next section, especially Proposition 2.7, and then go to Section 4 directly. From here to the end of the next section, we are dealing with a case which is a dual of the one studied in Section 4. However, for the reader who is unfamiliar with this topic, it is better just to follow the context since we present everything in detail in these two sections. A large part of the details in Sections 4 and 6 are omitted since they are supposed to be known.

To state the main results of this section, we need some notation. First, we define two operators as follows.

$$I_i(f) = \frac{1}{\mu_i b_i (f_i - f_{i+1})} \sum_{j \leqslant i} \mu_j f_j, \qquad I\!I_i(f) = \frac{1}{f_i} \sum_{j=i}^{N} \frac{1}{\mu_j b_j} \sum_{k \leqslant j} \mu_k f_k. \tag{2.14}$$

They are called an operator of single sum (integral) or double sum, respectively. Here for the first operator, we use a convention: $f_{N+1} = 0$ if $N < \infty$. The second operator can be alternatively expressed as

$$I\!I_i(f) = \frac{1}{f_i} \sum_{k \in E} \mu_k f_k \, \nu[i \vee k, N], \qquad \nu[\ell, m] = \sum_{i=\ell}^{m} \nu_i, \quad \nu_i = \frac{1}{\mu_i b_i}. \tag{2.15}$$

Next, define a difference operator R as follows.

$$R_i(v) = a_i\big(1 - v_{i-1}^{-1}\big) + b_i(1 - v_i), \quad i \in E, \ v_{-1} > 0 \text{ is free}, \ v_N := 0 \text{ if } N < \infty. \tag{2.16}$$

The domain of the operators $I\!I$, I and R are defined, respectively, as follows.

$$\mathscr{F}_{I\!I} = \{f : f > 0 \text{ on } E\},$$
$$\mathscr{F}_I = \{f : f > 0 \text{ on } E \text{ and is strictly decreasing}\},$$
$$\mathscr{V}_1 = \big\{v\colon \text{for all } i\ (0 \leqslant i < N),\ v_i \in (0,1) \text{ if } \textstyle\sum_j \nu_j < \infty \text{ and } v_i \in (0,1] \text{ if } \textstyle\sum_j \nu_j = \infty\big\}.$$

These sets are used for the lower estimates. For the upper estimates, we need some modifications of them as follows.

$$\widetilde{\mathscr{F}}_{I\!I} = \big\{f : f > 0 \text{ up to some } m : 1 \leqslant m < N+1 \text{ and then vanishes}\big\},$$
$$\begin{aligned}\widetilde{\mathscr{F}}_I = \big\{f : f &\text{ is strictly decreasing on some interval } [n, m]\ (0 \leqslant n < m < N+1),\\ &f_i = f_n \text{ for } i \leqslant n,\ f_m > 0, \text{ and } f_i = 0 \text{ for } i > m\big\},\end{aligned}$$
$$\begin{aligned}\widetilde{\mathscr{V}}_1 = \cup_{m=1}^{N-1}\big\{v : a_{i+1}&(a_{i+1} + b_{i+1})^{-1} < v_i < 1 - a_i\big(v_{i-1}^{-1} - 1\big)b_i^{-1}\\ &\text{for } i = 0, 1, \ldots, m-1 \text{ and } v_i = 0 \text{ for } i \geqslant m\big\}.\end{aligned}$$

Here and in what follows, to use the above operators on these modified sets, we adopt the usual convention $1/0 = \infty$. Besides, the operator II should be generalized as follows:

$$II_i(f) = \frac{1}{f_i} \sum_{i \leqslant j \in \text{supp}\,(f)} \frac{1}{\mu_j b_j} \sum_{k \leqslant j} \mu_k f_k, \qquad i \in \text{supp}\,(f). \tag{2.17}$$

From now on, we should remember that $II_\bullet(f)$ is defined on $\text{supp}\,(f)$ only. Fortunately, we need only to consider the following two cases: either $\text{supp}\,(f) = \{0, 1, \ldots, m\}$ for a finite m or $\text{supp}\,(f) = E$.

To avoid the heavy notation, we now split our main result of this section into a theorem and a proposition below.

Theorem 2.4. *The following variational formulas hold for λ_0 defined by (2.2).*

(1) *Difference form:*

$$\inf_{v \in \widetilde{\mathscr{V}}_1} \sup_{i \in E} R_i(v) = \lambda_0 = \sup_{v \in \mathscr{V}_1} \inf_{i \in E} R_i(v).$$

(2) *Single summation form:*

$$\inf_{f \in \widetilde{\mathscr{F}}_I} \sup_{i \in E} I_i(f)^{-1} = \lambda_0 = \sup_{f \in \mathscr{F}_I} \inf_{i \in E} I_i(f)^{-1}.$$

(3) *Double summation form:*

$$\inf_{f \in \widetilde{\mathscr{F}}_{II}} \sup_{i \in \text{supp}\,(f)} II_i(f)^{-1} = \lambda_0 = \sup_{f \in \mathscr{F}_{II}} \inf_{i \in E} II_i(f)^{-1}.$$

Moreover, the supremum on the right-hand side of the above three formulas can be attained.

The next result extends the domain of λ_0 or adds some additional sets of test functions for the operators I and II, respectively. Roughly speaking, a larger set of test functions provides more freedom in practice and a smaller one is helpful for producing a better estimate.

Proposition 2.5.

(1) *We have*

$$\lambda_0 = \inf\{D(f) : \|f\| = 1,\ f_{N+1} = 0\}, \tag{2.18}$$

where $f_\infty := \lim_{N \to \infty} f_N$ in the case of $N = \infty$.

(2) *When $\lambda_0 > 0$, the eigenfunction g satisfies $g_{N+1} = 0$.*

(3) *Moreover, we have*

$$\lambda_0 = \inf_{f \in \widetilde{\mathscr{F}}_I'} \sup_{i \in E} I_i(f)^{-1} \tag{2.19}$$

$$= \inf_{f \in \widetilde{\mathscr{F}}_{II} \cup \widetilde{\mathscr{F}}_{II}'} \sup_{i \in \text{supp}\,(f)} II_i(f)^{-1}, \tag{2.20}$$

$$\inf_{f \in \widetilde{\mathscr{F}}_I} \sup_{i \in \text{supp}\,(f)} II_i(f)^{-1} = \lambda_0 = \sup_{f \in \mathscr{F}_I} \inf_{i \in E} II_i(f)^{-1}, \tag{2.21}$$

where

$$\begin{aligned}\widetilde{\mathscr{F}}_I' &= \big\{f: f \textit{ is strictly decreasing and positive up to some } m: 1 \leqslant m < N+1\\ &\qquad \textit{and then vanishes}\big\} \subset \widetilde{\mathscr{F}}_I,\\ \widetilde{\mathscr{F}}_{II}' &= \big\{f: f>0 \textit{ on } E \textit{ and } f\,\mathit{II}(f)\in L^2(\mu)\big\}.\end{aligned}$$

Besides, the supremum $\sup_{f\in\mathscr{F}_I}$ *in (2.21) can also be attained.*

The condition "$f_{N+1}=0$" in (2.18) explains the meaning of "absorbing (Dirichlet) boundary at infinity" used in the title of this and the next sections.

Among the different groups of variational forms, the difference form is the simplest one in the practical computations. For instance, when $N=\infty$, by choosing $v_i \equiv c < 1$, we obtain the following simple lower estimate:

$$\lambda_0 \geqslant \inf_{i\in E}\,[b_i(1-c) - a_i(c^{-1}-1)].$$

This is non-trivial and is indeed sharp for a linear model (Example 3.5, $c=1/2$). The difference form of the variational formulas will be used in Section 5 to deduce a dual representation of λ_0. In general, the estimates produced by the operator R can be improved by using the operator I and further improved by using II. The price is that more computation is required successively. The single summation form of the variational formulas enables us to deduce a criterion for $\lambda_0>0$ (Theorem 3.1). Whereas the double summation form of the variational formulas enables us to deduce an approximating procedure to improve step by step the lower and upper estimates of λ_0 (Theorem 3.2).

Next, we mention that when $N=\infty$, for the upper estimates (the left-hand side of the formulas given in Theorem 2.4 or the formula given in (2.20)), the truncating procedure or the condition "$f\,\mathit{II}(f)\in L^2(\mu)$" cannot be removed. For instance, the formally dual formula $\inf_{0<v\leqslant 1}\sup_{i\in E} R_i(v)$ of the lower estimate $\sup_{0<v\leqslant 1}\inf_{i\in E} R_i(v)\,[=\sup_{v\in\mathscr{V}_1}\inf_{i\in E} R_i(v)]$ is not an upper bound of λ_0, and is indeed trivial. To see this, simply take $\bar v_i\equiv 1\,(i<\infty)$. Then $R_i(\bar v)\equiv 0$ and so

$$\inf_{0<v\leqslant 1}\sup_{i\in E} R_i(v) \leqslant \sup_{i\in E} R_i(\bar v) = 0.$$

More concretely, take $b_i\equiv 2$ and $a_i\equiv 1$. Then for $\bar v_i\equiv c<1$, we have

$$\inf_{v\in\mathscr{V}_1}\sup_{i\in E} R_i(v) \leqslant \inf_{c<1}\sup_{i\in E} R_i(\bar v) = 2\inf_{c<1}(1-c) = 0,$$

but $\lambda_0 = \big(\sqrt{2}-1\big)^2$ as will be seen in the next section (Example 3.4). Therefore, the quantity $\inf_{0<v\leqslant 1}\sup_{i\in E} R_i(v)$, as well as $\inf_{v>0}\sup_{i\in E} R_i(v)$, has no use for an upper estimate of λ_0.

Proofs of Theorem 2.4 *and Proposition* 2.5.

Part I. Recall that $\lambda_0^{(\#)}$ denotes the one defined by the formula (#). In particular, the notation λ_0 used from now on in this section is $\lambda_0^{(2.2)}$.

To prove the lower estimates, we adopt the following circle argument:

$$\begin{aligned}\lambda_0 &\geqslant \lambda_0^{(2.18)} \geqslant \sup_{f\in\mathscr{F}_{II}} \inf_{i\in E} \mathit{II}_i(f)^{-1} = \sup_{f\in\mathscr{F}_I} \inf_{i\in E} \mathit{II}_i(f)^{-1} = \sup_{f\in\mathscr{F}_I} \inf_{i\in E} I_i(f)^{-1}\\ &\geqslant \sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \geqslant \lambda_0. \end{aligned}\tag{2.22}$$

Clearly, $\lambda_0^{(2.18)} = \lambda_0$ if $N < \infty$. However, the identity is not trivial in the case of $N = \infty$. Besides, we will show that each supremum in (2.22) can be attained; and furthermore the eigenfunction g satisfies $g_{N+1} = 0$ whenever $\lambda_0 > 0$.

(a) Prove that $\lambda_0 \geqslant \lambda_0^{(2.18)} \geqslant \sup_{f\in\mathscr{F}_{II}} \inf_{i\in E} \mathit{II}_i(f)^{-1}$.

When $N = \infty$, the first inequality is trivial since

$$\{\|f\| = 1, f \in \mathscr{K}\} \subset \{\|f\| = 1,\ f_\infty = 0\}.$$

The proof of the second inequality is parallel to the first part of the proof of [4; Theorem 2.1]. Let g satisfy $g_{N+1} = 0$ and $\|g\| = 1$, and let (h_i) be a positive sequence. Then by a good use of the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}1 &= \sum_i \mu_i g_i^2 \quad (\text{since } \|g\| = 1)\\ &= \sum_i \mu_i \bigg(\sum_{j=i}^N (g_j - g_{j+1})\bigg)^2 \quad (\text{since } g_{N+1} = 0)\\ &\leqslant \sum_i \mu_i \sum_{j=i}^N \frac{(g_{j+1}-g_j)^2 \mu_j b_j}{h_j} \sum_{k=i}^N \frac{h_k}{\mu_k b_k}.\end{aligned}$$

Exchanging the order of the first two sums on the right-hand side, we get

$$\begin{aligned}1 &\leqslant \sum_j \mu_j b_j (g_{j+1}-g_j)^2 \frac{1}{h_j} \sum_{i\leqslant j} \mu_i \sum_{k=i}^N \frac{h_k}{\mu_k b_k}\\ &\leqslant D(g) \sup_{j\in E} \frac{1}{h_j} \sum_{i\leqslant j} \mu_i \sum_{k=i}^N \frac{h_k}{\mu_k b_k}\\ &=: D(g) \sup_{j\in E} H_j.\end{aligned}$$

We mention that the right-hand side may be infinite but we do not care at the moment. Now, let $f \in \mathscr{F}_{II}$ satisfy $c := \sup_{j\in E} \mathit{II}_j(f) < \infty$ and take $h_j = \sum_{i\leqslant j} \mu_i f_i$. Then $h_j \leqslant c f_j / v_j < \infty$ for all j. By the proportional property, we have

$$\sup_{j\in E} H_j \leqslant \sup_{j\in E} \frac{1}{f_j} \sum_{k=j}^N \frac{h_k}{\mu_k b_k} = \sup_{j\in E} \frac{1}{f_j} \sum_{k=j}^N \frac{1}{\mu_k b_k} \sum_{i\leqslant k} \mu_i f_i = \sup_{j\in E} \mathit{II}_j(f) < \infty.$$

Combining these facts together, we obtain $\lambda_0^{(2.18)} \geqslant \inf_{j\geqslant 0} I\!I_j(f)^{-1}$ whenever $\sup_{j\in E} I\!I_j(f) < \infty$. The inequality is trivial if $\sup_{j\in E} I\!I_j(f) = \infty$ and so it holds for all $f \in \mathscr{F}_{I\!I}$. By making the supremum with respect to $f \in \mathscr{F}_{I\!I}$, we obtain the required assertion.

(b) Prove that $\sup\limits_{f\in\mathscr{F}_{I\!I}} \inf\limits_{i\in E} I\!I_i(f)^{-1} = \sup\limits_{f\in\mathscr{F}_I} \inf\limits_{i\in E} I\!I_i(f)^{-1} = \sup\limits_{f\in\mathscr{F}_I} \inf\limits_{i\in E} I_i(f)^{-1}$.

Let $f \in \mathscr{F}_I \subset \mathscr{F}_{I\!I}$. Without loss of generality, assume that $\sup_{i\in E} I_i(f) < \infty$. By using the proportional property, we obtain

$$\begin{aligned}
\sup_{i\in E} I\!I_i(f) &= \sup_{i\in E} \frac{1}{f_i} \sum_{j=i}^{N} \frac{1}{\mu_j b_j} \sum_{k\leqslant j} \mu_k f_k \\
&\leqslant \sup_{i\in E} \sum_{j=i}^{N} \frac{1}{\mu_j b_j} \sum_{k\leqslant j} \mu_k f_k \bigg/ \sum_{j=i}^{N} (f_j - f_{j+1}) \quad (\text{since } f_{N+1} \geqslant 0) \\
&\leqslant \sup_{i\in E} \frac{1}{f_i - f_{i+1}} \bigg(\frac{1}{\mu_i b_i} \sum_{k\leqslant i} \mu_k f_k \bigg) \quad (\text{note that } f_i > f_{i+1}) \\
&= \sup_{i\in E} I_i(f) < \infty. \qquad (2.23)
\end{aligned}$$

Making the infimum with respect to $f \in \mathscr{F}_I$, we get

$$\inf_{f\in\mathscr{F}_I} \sup_{i\in E} I\!I_i(f) \leqslant \inf_{f\in\mathscr{F}_I} \sup_{i\in E} I_i(f).$$

Since $\mathscr{F}_I \subset \mathscr{F}_{I\!I}$, the left-hand side is bounded below by $\inf_{f\in\mathscr{F}_{I\!I}} \sup_{i\in E} I\!I_i(f)$. We have thus proved that

$$\sup_{f\in\mathscr{F}_{I\!I}} \inf_{i\in E} I\!I_i(f)^{-1} \geqslant \sup_{f\in\mathscr{F}_I} \inf_{i\in E} I\!I_i(f)^{-1} \geqslant \sup_{f\in\mathscr{F}_I} \inf_{i\in E} I_i(f)^{-1}.$$

There are two ways to prove the inverse inequality. The first one is longer but contains a useful technique. Let $f \in \mathscr{F}_{I\!I}$ with $c := \sup_{i\in E} I\!I_i(f) < \infty$. Set

$$g_i = \sum_{j=i}^{N} \frac{1}{\mu_j b_j} \sum_{k\leqslant j} \mu_k f_k = \sum_{j=i}^{N} \nu_j \sum_{k\leqslant j} \mu_k f_k > 0, \qquad i \in E,\ g_{N+1} := 0 \text{ if } N < \infty.$$

Then g_i is strictly decreasing in i, $g_i < g_0 \leqslant c f_0 < \infty$ for all i. Hence, $g \in \mathscr{F}_I$. Noticing that

$$g_i - g_{i+1} = \sum_{j=i}^{N} \nu_j \sum_{k\leqslant j} \mu_k f_k - \sum_{j=i+1}^{N} \nu_j \sum_{k\leqslant j} \mu_k f_k = \nu_i \sum_{k\leqslant i} \mu_k f_k$$

(here and in what follows, $\sum_{k=i}^{j}$ means $\sum_{i\leqslant k<j+1}$ and $\sum_{\emptyset} = 0$ by the standard convention), we have

$$\begin{aligned}
\Omega g(i) &= b_i(g_{i+1} - g_i) + a_i(g_{i-1} - g_i) \\
&= -b_i \nu_i \sum_{k\leqslant i} \mu_k f_k + a_i \nu_{i-1} \sum_{k\leqslant i-1} \mu_k f_k \\
&= -\frac{1}{\mu_i} \sum_{k\leqslant i} \mu_k f_k + \frac{a_i}{\mu_{i-1} b_{i-1}} \sum_{k\leqslant i-1} \mu_k f_k \\
&= -f_i, \qquad 1 \leqslant i < N.
\end{aligned}$$

Actually, this holds also for $i=0$ and $i=N$ if $N<\infty$. Applying (2.7) to $h=f$, by (2.9), it follows that

$$\mu_k b_k(g_k-g_{k+1})=\sum_{j\leqslant k}\mu_j g_j f_j/g_j \geqslant \sum_{j\leqslant k}\mu_j g_j \inf_{i\in E} I\!I_i(f)^{-1}, \qquad k\in E.$$

That is,

$$\sup_{i\in E} I\!I_i(f)\geqslant \frac{1}{\mu_k b_k(g_k-g_{k+1})}\sum_{j\leqslant k}\mu_j g_j = I_k(g), \qquad k\in E.$$

Making the supremum with respect to k, we obtain

$$\inf_{k\in E} I_k(g)^{-1}\geqslant \inf_{i\in E} I\!I_i(f)^{-1},$$

and hence,

$$\sup_{g\in\mathscr{F}_I}\inf_{k\in E} I_k(g)^{-1}\geqslant \inf_{i\in E} I\!I_i(f)^{-1}.$$

This lower bound becomes trivial if $\sup_{i\in E} I\!I_i(f)=\infty$, and hence, the inequality holds for all $f\in\mathscr{F}_{I\!I}$. Making the supremum with respect to $f\in\mathscr{F}_{I\!I}$, we obtain

$$\sup_{g\in\mathscr{F}_I}\inf_{k\in E} I_k(g)^{-1}\geqslant \sup_{f\in\mathscr{F}_{I\!I}}\inf_{i\in E} I\!I_i(f)^{-1}.$$

We have thus proved the required assertion.

The second proof is to show that

$$\sup_{f\in\mathscr{F}_I}\inf_{i\in E} I_i(f)^{-1}\geqslant \lambda_0$$

and thus completes a smaller circle argument. To do so, without loss of generality, assume that $\lambda_0>0$. Let $g>0$ be the eigenfunction of λ_0. Applying (2.9) to $h=\lambda_0 g$, we obtain $I_i(g)=\lambda_0^{-1}$ for all $i\in E$, and hence, $\inf_{i\in E} I_i(g)^{-1}=\lambda_0$. Noticing that $g\in\mathscr{F}_I$ by Proposition 2.2, the assertion is now obvious.

(c) Prove that $\sup_{f\in\mathscr{F}_{I\!I}}\inf_{i\in E} I\!I_i(f)^{-1}\geqslant \sup_{v\in\mathscr{V}_1}\inf_{i\in E} R_i(v)$.

Note that by a change of the sequence $\{v_i\}_{i=0}^{N-1}$:

$$u_i=v_0v_1\cdots v_{i-1}, \qquad i\in E,\ v_{-1}>0 \text{ is free},\ v_N:=0 \text{ if } N<\infty,$$

the quantity $R_i(v)$ becomes

$$a_i\bigg(1-\frac{u_{i-1}}{u_i}\bigg)+b_i\bigg(1-\frac{u_{i+1}}{u_i}\bigg), \qquad i\in E,\ u_{-1}>0 \text{ is free},\ u_{N+1}:=0 \text{ if } N<\infty.$$

To save our notation, we use $R_i(u)$ to denote this quantity. Clearly, $\{u_i\}$ is positive and $v_i\leqslant 1$ for all i mean that $\{u_i\}$ is non-increasing.

Before moving further, we prove that if $\inf_{i\in E} R_i(u) > 0$ for a positive sequence $u = (u_i)$, then u_i must be strictly decreasing in i. To do so, let

$$f_i = (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}.$$

Then $f_i = u_i R_i(u) > 0$ for all $i \in E$ by assumption, and so $f \in \mathscr{F}_{II}$. Noticing that

$$\mu_k f_k = \mu_{k+1} a_{k+1}(u_k - u_{k+1}) - \mu_k a_k (u_{k-1} - u_k),$$

we obtain

$$0 < \sum_{k\leqslant j} \mu_k f_k = \mu_{j+1} a_{j+1}(u_j - u_{j+1}) = \mu_j b_j (u_j - u_{j+1}).$$

Hence, u_i is strictly decreasing in i (equivalently, $v_i := u_{i+1}/u_i < 1$). This proves the required assertion. The reason of using $\mathscr{V}_1$ rather than $\{v : v_i > 0, 0 \leqslant i < N\}$ should be clear now.

We now return to our main assertion. For this, without loss of generality, assume that $\inf_{i\in E} R_i(u) > 0$ for a given strictly decreasing $u = (u_i)$. Otherwise, the assertion is trivial. From the last formula, we obtain

$$0 < \sum_{j=i}^{N} \nu_j \sum_{k\leqslant j} \mu_k f_k = \sum_{j=i}^{N} (u_j - u_{j+1}) = u_i - u_{N+1} \leqslant u_i.$$

Therefore,

$$0 < R_i(u) = \frac{f_i}{u_i} \leqslant f_i \bigg(\sum_{j=i}^{N} \nu_j \sum_{k\leqslant j} \mu_k f_k \bigg)^{-1} = II_i(f)^{-1}, \qquad i \in E.$$

It follows that

$$\inf_{i\in E} R_i(u) \leqslant \inf_{i\in E} II_i(f)^{-1} \leqslant \sup_{f\in\mathscr{F}_{II}} \inf_{i\in E} II_i(f)^{-1}.$$

The assertion now follows by making the supremum with respect to u.

(d) Prove that $\sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \geqslant \lambda_0$.

Assume that $\lambda_0 > 0$ for a moment (in particular, if $N < \infty$). Then by Proposition 2.2, the corresponding eigenfunction g (with $g_0 = 1$) of λ_0 is positive and strictly decreasing. From the eigenequation

$$-\Omega g(i) = \lambda_0 g_i, \qquad i \in E,\ g_{N+1} := 0 \text{ if } N < \infty,$$

it follows that

$$a_i \bigg(1 - \frac{g_{i-1}}{g_i}\bigg) + b_i \bigg(1 - \frac{g_{i+1}}{g_i}\bigg) = \lambda_0, \qquad i \in E.$$

Let $v_i = g_{i+1}/g_i$. Then $v_i \in (0,1)$ for all $i < N$ and so $v = (v_i) \in \mathscr{V}_1$. Moreover, $R_i(v) = \lambda_0$ for all $i \in E$. Therefore, we certainly have $\sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \geqslant \lambda_0$, as required.

It remains to prove that $\sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \geqslant 0$ when $N = \infty$. First, let $\sum_{i=0}^{\infty} \nu_i < \infty$. Choose a positive f such that

$$\sum_{k=0}^{\infty} \mu_k f_k \varphi_k < \infty, \qquad \varphi_k := \sum_{j=k}^{\infty} \nu_j.$$

Define

$$h_i = \sum_{j=i}^{\infty} \nu_j \sum_{k\leqslant j} \mu_k f_k, \qquad i \geqslant 0.$$

Then

$$h_i = \sum_{k=0}^{\infty} \mu_k f_k \varphi_{i\vee k} \leqslant \sum_{k=0}^{\infty} \mu_k f_k \varphi_k < \infty.$$

Set $\bar{\nu}_i = h_{i+1}/h_i \,(i \geqslant 0)$. Then $\bar{\nu} \in \mathscr{V}_1$ since h_i is strictly decreasing. A simple computation shows that $R_i(\bar{v}) = \mathit{II}_i(f)^{-1} > 0$ for all $i \geqslant 0$. Hence, $\sup_{v\in\mathscr{V}_1} \inf_{i\geqslant 0} R_i(v) \geqslant 0$. Next, let $\sum_i \nu_i = \infty$ and set $\bar{v}_i \equiv 1$. Then $R_i(\bar{v}) \equiv 0$ and so the same conclusion holds.

The proof of the last paragraph indicates the reason why in $\mathscr{V}_1$ we define "$v_i \in (0,1)$" and "$v_i \in (0,1]$" separately according to "$\sum_i \nu_i < \infty$" or "$\sum_i \nu_i = \infty$". Although we have known from proof (c) that for $\inf_i R_i(v) > 0$, it is necessary that $v < 1$ but this condition may not be sufficient for $\inf_i R_i(v) \geqslant 0$. The extremal $\bar{v}_i \equiv 1$ is used only in the case of $\sum_i \nu_i = \infty$ in which we indeed have $\lambda_0 = 0$ (cf. Theorem 3.1 below).

We have thus completed the proof of circle (2.22).

(e) We now prove that each supremum in (2.22) can be attained. The case that $\lambda_0 = 0$ is easier since

$$0 = \lambda_0 \geqslant \inf_{i\in E} \mathit{II}_i(f)^{-1} \geqslant 0 \quad \text{and} \quad 0 = \lambda_0 \geqslant \inf_{i\in E} I_i(f)^{-1} \geqslant 0$$

for every f in the corresponding domain, as an application of (2.22). Similarly, the conclusion holds for the operator R as seen from proof (d): noting that in the degenerated case that $\sum_i \nu_i = \infty$, we have $\lambda_0 = 0$ and then $v_i \equiv 1$ by Proposition 2.2 (3).

Next, we consider the case that $\lambda_0 > 0$ with eigenfunction g: $g_0 = 1$. Then for the operator R, the supremum is attained at $v_i = g_{i+1}/g_i$ as seen from the first paragraph of proof (d). For the operator I, it is attained at $f = g$ as an application of Proposition 2.1 with $c_i \equiv 0$: $I_i(g) \equiv \lambda_0^{-1}$. At the same time, in view of part (2) of Proposition 2.2, we have $\mathit{II}_i(g) \equiv \lambda_0^{-1}$ whenever $g_{N+1} = 0$.

It remains to rule out the possibility that $g_{N+1} > 0$. Otherwise, by part (2) of Proposition 2.2 again, we have $N = \infty$ and

$$M_i := \sum_{j\geqslant i} \nu_j \sum_{k\leqslant j} \mu_k \in (0,\infty).$$

Let $\tilde{g} = g - g_\infty$. Then $\tilde{g} \in \mathscr{F}_{II}$. Noting that

$$\begin{aligned}\sum_{j\geqslant i}\nu_j\sum_{k\leqslant j}\mu_k\tilde{g}_k &= \sum_{j\geqslant i}\nu_j\sum_{k\leqslant j}\mu_k g_k - g_\infty M_i\\ &= \frac{g_i - g_\infty}{\lambda_0} - g_\infty M_i \quad \text{(by (2.11))},\end{aligned}$$

we obtain

$$\sup_{i\geqslant 0} II_i(\tilde{g}) = \sup_{i\geqslant 0}\left[\frac{1}{\lambda_0} - \frac{g_\infty M_i}{g_i - g_\infty}\right] = \frac{1}{\lambda_0} - g_\infty \inf_{i\geqslant 0}\frac{M_i}{g_i - g_\infty}.$$

By using the proportional property and (2.5), it follows that

$$\inf_{i\geqslant 0}\frac{M_i}{g_i - g_\infty} \geqslant \inf_{i\geqslant 0}\frac{\nu_i\sum_{k\leqslant i}\mu_k}{g_i - g_{i+1}} = \frac{1}{\lambda_0}.$$

Thus, we get

$$\sup_{i\geqslant 0} II_i(\tilde{g}) \leqslant \frac{1}{\lambda_0}(1 - g_\infty) < \frac{1}{\lambda_0}.$$

Hence, $\inf\limits_{i\geqslant 0} II_i(\tilde{g})^{-1} > \lambda_0$, which is a contradiction to proof (a): $\lambda_0 \geqslant \inf\limits_{i\in E} II(\tilde{g})^{-1}$. We have thus proved that $g_\infty = 0$ whenever $\lambda_0 > 0$. Note that this paragraph uses Proposition 2.2 and proof (a) only.

Part II. Next, to prove the upper estimates, we adopt the following circle argument:

$$\begin{aligned}\lambda_0 &\leqslant \inf_{f\in\widetilde{\mathscr{F}}_{II}\cup\widetilde{\mathscr{F}}'_{II}}\ \sup_{i\in\text{supp}\,(f)} II_i(f)^{-1} && (2.24)\\ &\leqslant \inf_{f\in\widetilde{\mathscr{F}}_{II}}\ \sup_{i\in\text{supp}\,(f)} II_i(f)^{-1} && (2.25)\\ &= \inf_{f\in\widetilde{\mathscr{F}}_{I}}\ \sup_{i\in\text{supp}\,(f)} II_i(f)^{-1} = \inf_{f\in\widetilde{\mathscr{F}}_{I}}\ \sup_{i\in E} I_i(f)^{-1} && (2.26)\\ &\leqslant \inf_{f\in\widetilde{\mathscr{F}}'_{I}}\ \sup_{i\in E} I_i(f)^{-1} && (2.27)\\ &\leqslant \inf_{v\in\widetilde{\mathscr{V}}_1}\ \sup_{i\in E} R_i(v) && (2.28)\\ &\leqslant \lambda_0. && (2.29)\end{aligned}$$

Since inequalities (2.25) and (2.27) are obvious, we need only to prove (2.24), (2.26), (2.28) and (2.29).

(f) Prove that $\lambda_0 \leqslant \inf_{f\in\widetilde{\mathscr{F}}_{II}\cup\widetilde{\mathscr{F}}'_{II}} \sup_{i\in\text{supp}\,(f)} II_i(f)^{-1}$.

We remark that in the particular case that the eigenfunction f is in $L^2(\mu)$, then the function $g := f II(f)$ is nothing but just $f/\lambda_0 \in L^2(\mu)$. Hence, the infimum in (2.24) is attained at this $f \in \widetilde{\mathscr{F}'_{II}}$ and the equality sign in (2.24) holds.

We now consider the general case. Let $f \in \widetilde{\mathscr{F}}_{II}$. Then there exists an m such that $f_i > 0$ for $i \leqslant m$ and $f_i = 0$ for $i > m$. Set $g = \mathbb{1}_{\text{supp}\,(f)} f I\!I(f)$. That is,

$$g_i = \begin{cases} \sum_{j=i}^m \nu_j \sum_{k\leqslant j} \mu_k f_k, & i \leqslant m \\ 0, & i \geqslant m+1. \end{cases}$$

Clearly, $g \in L^2(\mu)$ and

$$g_i - g_{i+1} = \begin{cases} \nu_i \sum_{k\leqslant i} \mu_k f_k, & i \leqslant m \\ 0, & i \geqslant m+1. \end{cases}$$

We now have

$$D(g) = \sum_{i\leqslant m} \mu_i b_i (g_{i+1} - g_i)^2 = \sum_{i\leqslant m} (g_i - g_{i+1}) \sum_{k\leqslant i} \mu_k f_k = \sum_{k\leqslant m} \mu_k f_k \sum_{k\leqslant i\leqslant m} (g_i - g_{i+1}).$$

Since $g_{m+1} = 0$, we get

$$D(g) = \sum_{k\leqslant m} \mu_k f_k g_k \leqslant \sum_{k\leqslant m} \mu_k g_k^2 \max_{0\leqslant i\leqslant m} (f_i/g_i) = \|g\|^2 \sup_{i\in\text{supp}\,(f)} I\!I_i(f)^{-1}.$$

Dividing both sides by $\|g\|^2 \in (0,\infty)$, it follows that

$$\lambda_0 \leqslant D(g)/\|g\|^2 \leqslant \sup_{i\in\text{supp}\,(f)} I\!I_i(f)^{-1}, \qquad f \in \widetilde{\mathscr{F}}_{II}. \tag{2.30}$$

For $f \in \widetilde{\mathscr{F}}_{II}'$, the same conclusion clearly holds if $N < \infty$. When $N = \infty$, since $g \in L^2(\mu)$ by assumption, we have $0 < g < \infty$. As a tail sequence of a convergent series (which sum equals g_0), we certainly have $g_i \downarrow g_\infty = 0$ as $i \uparrow \infty$. Hence, the same proof replacing m with ∞, plus the fact that $\lambda_0 = \lambda_0^{(2.18)}$ proved in Part I, shows that

$$\lambda_0 = \lambda_0^{(2.18)} \leqslant \sup_{i\in\text{supp}\,(f)} I\!I_i(f)^{-1}, \qquad f \in \widetilde{\mathscr{F}}_{II}'.$$

Combining this with (2.30), we prove the required assertion.

The proof indicates the reason why the truncating procedure is used for the upper estimates since in general the eigenfunction g may not belong to $L^2(\mu)$ as shown by Proposition 2.2.

(g) Prove that

$$\inf_{f\in\widetilde{\mathscr{F}}_{II}} \sup_{i\in\text{supp}\,(f)} I\!I_i(f)^{-1} = \inf_{f\in\widetilde{\mathscr{F}}_{I}} \sup_{i\in\text{supp}\,(f)} I\!I_i(f)^{-1} = \inf_{f\in\widetilde{\mathscr{F}}_{I}} \sup_{i\in E} I_i(f)^{-1}.$$

Let $f \in \widetilde{\mathscr{F}}_I$. Then there exist $n < m$ such that $f_i = f_{i\vee n} \mathbb{1}_{\{i\leqslant m\}}$, $f_m > 0$, and f is strictly decreasing on $[n, m]$. Clearly, we have

$$\min_{i\leqslant m} I\!I_i(f) = \min_{n\leqslant i\leqslant m} I\!I_i(f) \quad \text{and} \quad \inf_{i\in E} I_i(f) = \min_{n\leqslant i\leqslant m} I_i(f)$$

since, by assumption, $1/0=\infty$. By the proportional property, first we have

$$\begin{aligned}\min_{n\leqslant i\leqslant m} I\!I_i(f) &= \min_{n\leqslant i\leqslant m} \sum_{j=i}^m \nu_j \sum_{k\leqslant j}\mu_k f_k \Big/ \sum_{j=i}^m (f_j-f_{j+1})\\ &\geqslant \min_{n\leqslant i\leqslant m} \frac{1}{\mu_i b_i (f_i-f_{i+1})}\sum_{k\leqslant i}\mu_k f_k\\ &= \min_{n\leqslant i\leqslant m} I_i(f),\end{aligned}$$

and then

$$\sup_{f\in\widetilde{\mathscr{F}}_{I\!I}}\ \inf_{i\in\text{supp}\,(f)} I\!I_i(f) \geqslant \sup_{f\in\widetilde{\mathscr{F}}_{I}}\ \inf_{i\in\text{supp}\,(f)} I\!I_i(f) \geqslant \sup_{f\in\widetilde{\mathscr{F}}_{I}}\ \inf_{i\in E} I_i(f)$$

since $\widetilde{\mathscr{F}}_I \subset \widetilde{\mathscr{F}}_{I\!I}$.

As in proof (b), there are two ways to prove the inverse inequality. First, let $f\in\widetilde{\mathscr{F}}_{I\!I}$. As in proof (f), set $g=\mathbb{1}_{\text{supp}\,(f)} f I\!I(f)$. Clearly, $g\in\widetilde{\mathscr{F}}_I' \subset \widetilde{\mathscr{F}}_I$ and moreover,

$$\begin{aligned}b_i(g_{i+1}-g_i)+a_i(g_{i-1}-g_i) &= -\frac{1}{\mu_i}\sum_{k\leqslant i}\mu_k f_k + \frac{a_i}{\mu_{i-1}b_{i-1}}\sum_{k\leqslant i-1}\mu_k f_k\\ &= -\frac{1}{\mu_i}\sum_{k\leqslant i}\mu_k f_k + \frac{1}{\mu_i}\sum_{k\leqslant i-1}\mu_k f_k\\ &= -f_i, \qquad i\leqslant m.\end{aligned}$$

When $i=0$, the second term on the left-hand side disappears since $a_0=0$. It follows that

$$\mu_i b_i(g_{i+1}-g_i)+\mu_i a_i(g_{i-1}-g_i) = -\mu_i f_i, \qquad i\leqslant m,$$

and furthermore,

$$\mu_k b_k(g_k-g_{k+1}) = \sum_{j\leqslant k}\mu_j g_j f_j/g_j \leqslant \sum_{j\leqslant k}\mu_j g_j \max_{0\leqslant i\leqslant m} I\!I_i(f)^{-1}, \qquad k\leqslant m.$$

That is,

$$\min_{0\leqslant i\leqslant m} I\!I_i(f) \leqslant \frac{1}{\mu_k b_k(g_k-g_{k+1})}\sum_{j\leqslant k}\mu_j g_j = I_k(g), \qquad k\leqslant m.$$

Making the infimum with respect to k, we obtain

$$\max_{0\leqslant k\leqslant m} I_k(g)^{-1} \leqslant \max_{0\leqslant i\leqslant m} I\!I_i(f)^{-1}.$$

One may rewrite $\max_{0\leqslant k\leqslant m}$ as $\sup_{k\in E}$ on the left-hand side since $I_k(g)=\infty$ for all $k\geqslant m+1$. Since $g\in\widetilde{\mathscr{F}}_I'\subset\widetilde{\mathscr{F}}_I$, we now have

$$\inf_{g\in\widetilde{\mathscr{F}}_I}\sup_{k\in E}I_k(g)^{-1}\leqslant\inf_{g\in\widetilde{\mathscr{F}}_I'}\sup_{k\in E}I_k(g)^{-1}\leqslant\sup_{i\in\operatorname{supp}(f)}\mathit{II}_i(f)^{-1}.$$

Next, making the infimum with respect to $f\in\widetilde{\mathscr{F}}_{II}$, we obtain

$$\inf_{g\in\widetilde{\mathscr{F}}_I}\sup_{k\in E}I_k(g)^{-1}\leqslant\inf_{g\in\widetilde{\mathscr{F}}_I'}\sup_{k\in E}I_k(g)^{-1}\leqslant\inf_{f\in\widetilde{\mathscr{F}}_{II}}\sup_{i\in\operatorname{supp}(f)}\mathit{II}_i(f)^{-1}.$$

The second proof for the inverse inequality is to show that

$$\inf_{f\in\widetilde{\mathscr{F}}_I'}\sup_{i\in E}I_i(f)^{-1}\leqslant\lambda_0.$$

For this, recall the definition

$$\lambda_0=\inf\{D(f):\|f\|=1,\ f_i=0\text{ for all }i>\text{some }m:1\leqslant m<N+1\}.$$

Because of

$$\begin{aligned}&\{\|f\|=1,f_i=0\text{ for all }i>m:1\leqslant m<N+1\}\\&\quad\subset\{\|f\|=1,\ f_i=0\text{ for all }i>m+1:1\leqslant m<N+1\},\end{aligned}$$

it is clear that

$$\lambda_0^{(m)}:=\inf\{D(f):\|f\|=1,\ f_i=0\text{ for all }i>m:1\leqslant m<N+1\}\ \downarrow\lambda_0$$

as $m\uparrow N$. Note that $\lambda_0^{(m)}$ is just the first eigenvalue of the Dirichlet form $(D,\mathscr{D}(D))$ restricted to $\{0,1,\ldots,m\}$ with Dirichlet (absorbing) boundary at $m+1$. Now, let $g=g^{(m)}$ be the eigenfunction of $\lambda_0^{(m)}>0$ with $g_0=1$. Extend g to the whole space by setting $g_i=0$ for all $i>m$. By using Proposition 2.2, it follows that $g\in\widetilde{\mathscr{F}}_I'$ with $\operatorname{supp}(g)=\{0,1,\ldots,m\}$. Furthermore, by (2.9) with $h=\lambda_0 g$, we have $I_i(g)^{-1}=\lambda_0^{(m)}>0$ for all $i\leqslant m$, and hence,

$$\sup_{i\in E}I_i(g)^{-1}=\sup_{i\leqslant m}I_i(g)^{-1}=\lambda_0^{(m)}.$$

Thus,

$$\lambda_0^{(m)}=\sup_{i\in E}I_i(g)^{-1}\geqslant\inf_{f\in\widetilde{\mathscr{F}}_I',\ \operatorname{supp}(f)=\{0,1,\ldots,m\}}\sup_{i\in E}I_i(f)^{-1}\geqslant\inf_{f\in\widetilde{\mathscr{F}}_I'}\sup_{i\in E}I_i(f)^{-1}.$$

The assertion now follows by letting $m\to N$.

(h) Prove that $\inf_{f\in\widetilde{\mathscr{F}}_{II}}\sup_{i\in\operatorname{supp}(f)}\mathit{II}_i(f)^{-1}\leqslant\inf_{v\in\widetilde{\mathscr{V}}_1}\sup_{i\in E}R_i(v)$.

Let u with $\operatorname{supp}(u) = \{0, 1, \dots, m\}$ be given such that $v_i := u_{i+1}/u_i \in \widetilde{\mathscr{V}}_1$. Then, the constraint

$$v_i < 1 - a_i\big(v_{i-1}^{-1} - 1\big)b_i^{-1}, \qquad 0 \leqslant i \leqslant m,\ v_m = 0,$$

is equivalent to $\min_{0\leqslant i\leqslant m} R_i(v) > 0$, and the constraint

$$v_i > a_{i+1}(a_{i+1} + b_{i+1})^{-1}, \qquad 0 \leqslant i \leqslant m-1,$$

comes from the requirement that $v_i > 0$ for all $i < m$. Since the case of $i = m$ in the first constraint is contained in the second one, we obtain the constraint described in $\widetilde{\mathscr{V}}_1$. In particular, we have

$$a_1(a_1 + b_1)^{-1} < v_0 < 1 - a_0\big(v_{-1}^{-1} - 1\big) = 1$$

and so $v_0 \in (0, 1)$. By induction, we have $v_i \in (0, 1)$ for all $i < m$. The existence of such a u is guaranteed since $m < \infty$, as will be shown in proof (i) below. Now, let

$$f_i = \begin{cases} (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}, & i \leqslant m, \\ 0, & i > m. \end{cases}$$

Then by assumption, $f_i/u_i = R_i(u) > 0$ for $i \leqslant m$. Hence, $f \in \widetilde{\mathscr{F}}_{II}$. Next, we have

$$0 < \sum_{k\leqslant j} \mu_k f_k = \mu_j b_j(u_j - u_{j+1}), \qquad j \leqslant m.$$

Hence,

$$\sum_{j=i}^{m} \nu_j \sum_{k\leqslant j} \mu_k f_k = u_i - u_{m+1} = u_i > 0, \qquad i \leqslant m.$$

Therefore, we obtain

$$R_i(u) = \frac{f_i}{u_i} = f_i \Big/ \sum_{j=i}^{m} \nu_j \sum_{k\leqslant j} \mu_k f_k = II_i(f)^{-1}, \qquad i \leqslant m$$

and then

$$\sup_{i\in E} R_i(u) = \max_{i\leqslant m} R_i(u) = \sup_{i\in \operatorname{supp}(f)} II_i(f)^{-1} \geqslant \inf_{f\in\widetilde{\mathscr{F}}_{II}} \sup_{i\in \operatorname{supp}(f)} II_i(f)^{-1}.$$

To be consistent with the convention of $R_i(v)$, here we adopt the convention: $R_i(u) = -\infty$ for all $i > m$. The assertion now follows by making the infimum with respect to u.

(i) Prove that $\inf_{v\in\widetilde{\mathscr{V}}_1} \sup_{i\in E} R_i(v) \leqslant \lambda_0$.

As in the last part of proof (g), denote by g (with $g_0 = 1$) the eigenfunction of $\lambda_0^{(m)} > 0$. Then $\operatorname{supp}(g) = \{0, 1, \dots, m\}$, and g is strictly decreasing on $\{0, 1, \dots, m\}$ by part (2) of Proposition 2.2. The definition of g gives us

$$b_i(g_i - g_{i+1}) - a_i(g_{i-1} - g_i) = \lambda_0^{(m)} g_i, \qquad i \leqslant m,\ g_{m+1} = 0.$$

That is,

$$a_i\left(1-\frac{g_{i-1}}{g_i}\right)+b_i\left(1-\frac{g_{i+1}}{g_i}\right)=\lambda_0^{(m)},\qquad i\leqslant m.$$

Let $v_i=g_{i+1}/g_i$ for $i\leqslant m$ and $v_i=0$ for $i>m$. Then $v_i\in(0,1)$ for $i\in\{0,1,\ldots,m-1\}$, and $R_i(v)=\lambda_0^{(m)}$ for all $i\leqslant m$. It is now easy to see that $v\in\widetilde{\mathscr{V}}_1$. We have thus constructed a $u\,(=g)$ required in proof (h). Clearly $R_i(v)=-\infty$ for all $i>m$. Therefore,

$$\begin{aligned}\lambda_0^{(m)}&=\max_{0\leqslant i\leqslant m}R_i(v)\\&\geqslant\inf_{v\in\widetilde{\mathscr{V}}_1:\,\mathrm{supp}\,(v)=\{0,1,\ldots,m-1\}}\ \max_{0\leqslant i\leqslant m}R_i(v)\\&\geqslant\inf_{v\in\widetilde{\mathscr{V}}_1:\,\mathrm{supp}\,(u)=\{0,1,\ldots,n\}\text{ for some }n\geqslant 0}\ \sup_{i\in E}R_i(u)\\&=\inf_{v\in\widetilde{\mathscr{V}}_1}\sup_{i\in E}R_i(v).\end{aligned}$$

Letting $m\to N$, we obtain the required assertion.

We have thus completed the circle argument of (2.24)–(2.29) and then the proofs of Theorem 2.4 and Proposition 2.5 are finished. $\square$

Before moving further, we mention a technical point in the proof above. Instead of the approximation with finite state space used in Part II of the above proof, it seems more natural to use the truncating procedure for the eigenfunction g. However, the next result shows that this procedure is not practical in general.

Remark 2.6. *Let $g\neq 0$ be the eigenfunction of $\lambda_0>0$ and define $g^{(m)}=g\mathbb{1}_{\leqslant m}$. Then*

$$\min_{i\,\in\,\mathrm{supp}\,(g^{(m)})}I\!I_i\big(g^{(m)}\big)=\frac{1}{\lambda_0}\left[1-\frac{g_{m+1}}{g_m}\right].$$

In particular, the sequence $\big\{\min_{i\in\mathrm{supp}\,(g^{(m)})}I\!I_i\big(g^{(m)}\big)\big\}_{m\geqslant 1}$ may not converge to λ_0^{-1} as $m\uparrow\infty$.

Proof. Note that

$$\begin{aligned}\min_{i\in\mathrm{supp}\,(g^{(m)})}I\!I_i\big(g^{(m)}\big)&=\min_{0\leqslant i\leqslant m}\frac{1}{g_i}\sum_{j=i}^{m}\nu_j\sum_{k\leqslant j}\mu_k g_k\\&=\min_{0\leqslant i\leqslant m}\frac{1}{g_i}\sum_{j=i}^{m}\nu_j\frac{\mu_j b_j(g_j-g_{j+1})}{\lambda_0}\qquad(\text{by }(2.5))\\&=\min_{0\leqslant i\leqslant m}\frac{1}{\lambda_0 g_i}(g_i-g_{m+1})\\&=\frac{1}{\lambda_0}\left[1-\frac{g_{m+1}}{g_m}\right].\end{aligned}$$

This proves the main assertion. For Example 3.4 in the next section, we have

$$\lim_{m\to\infty}\left(1-\frac{g_{m+1}}{g_m}\right)=1-\sqrt{\frac{a}{b}}<1,$$

and so

$$\lim_{m\to\infty} \min_{i \in \operatorname{supp}(g^{(m)})} II_i\big(g^{(m)}\big) < {\lambda_0}^{-1}. \quad \square$$

To conclude this section and also for later use, we introduce a variational formula of λ_0 in a different difference form.

Proposition 2.7. *On the set* $\mathscr{V} := \{v : v_i > 0,\ 0 \leqslant i < N\}$*, redefine*

$$R_i(v) = a_{i+1} + b_i - a_i/v_{i-1} - b_{i+1}v_i, \qquad i \in E,\ v_{-1} > 0 \textit{ is free},$$

where $a_{N+1} = b_{N+1} = 0$ *and* v_N *is free if* $N < \infty$*. Then*

(1) *we have*

$$\sup_{v\in\mathscr{V}} \inf_{i\in E} R_i(v) \geqslant \lambda_0. \tag{2.31}$$

The equality sign holds once $\sum_{i=0}^N \mu_i = \infty$*. In this case, we indeed have*

$$\lambda_0 = \sup_{v\in\mathscr{V}} \inf_{i\in E} R_i(v) = \sup_{v\in\mathscr{V}_*} \inf_{i\in E} R_i(v),$$

where

$$\mathscr{V}_* = \{v : v_{i-1} > a_i/b_i,\ 0 \leqslant i < N+1\}.$$

(2) *In general, we have*

$$\lambda_0 = \sup_{v\in\mathscr{V}_*} \inf_{i\in E} R_i(v), \tag{2.32}$$

and the supremum in (2.32) can be attained.

Proof. (a) First, we prove that $\sup_{v\in\mathscr{V}_*} \inf_{i\in E} R_i(v) \geqslant 0$. Given a positive, non-increasing f, $f_{N+1} = 0$ if $N < \infty$, define

$$u_i = (\mu_i b_i)^{-1} \sum_{j\leqslant i} \mu_j f_j \in (0,\infty), \qquad i < N+1.$$

Then

$$b_i u_i - a_i u_{i-1} = f_i > 0, \qquad i \in E,\ u_{-1} > 0 \text{ is free}.$$

This implies that $(v_i := u_{i+1}/u_i : i < N) \in \mathscr{V}_*$. As before, we also use

$$R_i(u) := a_{i+1} + b_i - a_i \frac{u_{i-1}}{u_i} - b_{i+1}\frac{u_{i+1}}{u_i}, \qquad i \in E$$

instead of $R_i(v)$. Clearly,

$$R_i(u) = \frac{f_i - f_{i+1}}{u_i} \geqslant 0, \qquad i \in E.$$

Hence $\inf_{i\in E} R_i(u) \geqslant 0$ and the required assertion is now obvious.

(b) By (a), without loss of generality, assume that $\lambda_0 > 0$. Then by Proposition 2.2, the corresponding eigenfunction g of λ_0 is positive and strictly decreasing. With $u_i := g_i - g_{i+1} > 0\,(i \in E)$, the eigenequation

$$-\Omega g(i) = b_i u_i - a_i u_{i-1} = \lambda_0 g_i, \qquad i \in E, \quad g_{N+1} := 0 \text{ if } N < \infty,$$

gives us $(v_i := u_{i+1}/u_i : i < N) \in \mathscr{V}_*$. Next, by making a difference of $-\Omega g(i)$ and $-\Omega g(i+1)$ and noting that $\Omega g(N+1)$ is setting to be zero if $N < \infty$, we obtain

$$(a_{i+1} + b_i)u_i - b_{i+1}u_{i+1} - a_i u_{i-1} = \lambda_0 u_i, \qquad i \in E.$$

Thus, we have $R_i(v) = \lambda_0$ for all $i \in E$. Therefore, (2.31) holds.

(c) To prove the equality sign in (2.31) whenever $\sum_{i=0}^N \mu_i = \infty$, in view of Part I of the proofs of Theorem 2.4 and Proposition 2.5 and (b), it suffices to show that

$$\sup_{f \in \mathscr{F}_I} \inf_{i \in E} I_i(f)^{-1} \geqslant \sup_{v \in \mathscr{V}} \inf_{i \in E} R_i(v).$$

In view of (a), without loss of generality, assume that $\inf_{i \in E} R_i(u) > 0$ for a given $u > 0$. Define $f_i = b_i u_i - a_i u_{i-1}$ for $i \in E$, $f_{N+1} = 0$ of $N < \infty$. Then it is clear that

$$(f_i - f_{i+1})/u_i = R_i(u) > 0, \qquad i \in E. \tag{2.33}$$

Hence, f is strictly decreasing.

We now prove that $f \in \mathscr{F}_I$ whenever $\sum_i \mu_i = \infty$. First, we have

$$\sum_{k \leqslant i} \mu_k f_k = \sum_{k \leqslant i} \mu_k (b_k u_k - a_k u_{k-1}) = \sum_{k \leqslant i} (\mu_k b_k u_k - \mu_{k-1} b_{k-1} u_{k-1}) = \mu_i b_i u_i > 0, \quad i \in E. \tag{2.34}$$

In particular, $f_0 > 0$. If $f_{k_0} \leqslant 0$ for some $k_0 \geqslant 1$, then $f_{k_0+1} < 0$ and

$$\sum_{k_0+1 \leqslant i \leqslant n} \mu_i f_i < f_{k_0+1} \sum_{k_0 \leqslant i \leqslant n} \mu_i \to -\infty \qquad \text{as } n \to \infty$$

since $\sum_{i=0}^N \mu_i = \infty$. This implies that

$$\sum_{k_0+1 \leqslant i \leqslant n} \mu_i f_i \to -\infty \qquad \text{as } n \to \infty.$$

Now, by (2.33), we would get

$$0 < \mu_n b_n u_n = \sum_{i \leqslant n} \mu_i f_i = \sum_{i \leqslant k_0} \mu_i f_i + \sum_{k_0+1 \leqslant i \leqslant n} \mu_i f_i \to -\infty \qquad \text{as } n \to \infty,$$

which is impossible. Therefore, $f > 0$ and then $f \in \mathscr{F}_I$.

Combining (2.33) with (2.34), we obtain that

$$R_i(u) = \frac{f_i - f_{i+1}}{u_i} = \mu_i b_i (f_i - f_{i+1}) \Big/ \sum_{k \leqslant i} \mu_k f_k = I_i(f)^{-1}, \qquad i \in E.$$

Hence, we have first

$$\inf_{i\in E} R_i(u) = \inf_{i\in E} I_i(f)^{-1} \leqslant \sup_{f\in\mathscr{F}_I}\inf_{i\in E} I_i(f)^{-1},$$

and then

$$\sup_{u>0}\inf_{i\in E} R_i(u) \leqslant \sup_{f\in\mathscr{F}_I}\inf_{i\in E} I_i(f)^{-1},$$

as required. We have thus proved the equality in (2.31) under $\sum_i \mu_i = \infty$.

Actually, we have proved in the last paragraph that $(f_i =) b_i u_i - a_i u_{i-1} > 0$ for all $i \in E$ and so $(v_i := u_{i+1}/u_i) \in \mathscr{V}_*$ whenever $\inf_{i\in E} R_i(u) > 0$. This means that the set $\mathscr{V} \setminus \mathscr{V}_*$ is useless since for each $v \in \mathscr{V}\setminus\mathscr{V}_*$, we have $\inf_{i\in E} R_i(u) \leqslant 0$. Now, because of $\mathscr{V}_* \subset \mathscr{V}$ and (a), using the equality in (2.31), we obtain the last assertion of part (1).

(d) To prove part (2) of the proposition, note that the inequality "$\leqslant$" is proved in (b). For the inverse inequality, recalling that the main body in proof (c) is to show that the function f_i $(i \in E)$ defined there is positive, this is now automatic due to the definition of $\mathscr{V}_*$. The equality sign in (2.32) has already checked in proofs (a) and (b) in the cases $\lambda_0 = 0$ and $\lambda_0 > 0$, respectively. $\square$

Remark 2.8. For the equality in (2.31), the condition $\sum_i \mu_i = \infty$ cannot be removed. For instance, consider the ergodic case for which $\sum_i \mu_i < \infty$ but $\lambda_0 = 0$ by Theorem 3.1 below and so (2.31) is trivial. However, as proved in [3; Theorem 1.1] (cf. Theorem 6.1 below), the left-hand side of (2.31) coincides with another eigenvalue (called λ_1) which can be positive. In this case, the equality in (2.31) fails. This also explains the reason for the use of $\mathscr{V}_*$.

Remark 2.9. The test sequences with the same notation (v_i) used in Theorem 2.4 and Proposition 2.7 are usually different. Corresponding to the eigenfunction (g_i) of λ_0, the sequence constructed in proof (d) of Theorem 2.4 and Proposition 2.5 is $v_i = g_{i+1}/g_i$, but the one constructed in proof (b) of Proposition 2.7 is

$$v_i = \frac{g_{i+1} - g_{i+2}}{g_i - g_{i+1}} = \frac{1 - g_{i+2}/g_{i+1}}{g_i/g_{i+1} - 1}.$$

Thus, the mapping from the first sequence to the second one is as follows:

$$(v_i)_{0\leqslant i<N} \to \left(\frac{1 - v_{i+1}}{v_i^{-1} - 1}\right)_{0\leqslant i<N}, \tag{2.35}$$

where on the right-hand side, v_N is set to be zero if $N < \infty$.

3. Absorbing (Dirichlet) boundary at infinity: criterion, approximating procedure and examples

This section is a continuation of the last one. As applications of the variational formulas given in the last section, a criterion for the positivity of λ_0 and an approximating procedure for λ_0 are presented. The section is ended by a class of examples and then the study on the first case of our classification is completed.

Theorem 3.1 (Criterion and basic estimates). *The decay rate* $\lambda_0 > 0$ *iff* $\delta < \infty$, *where*

$$\delta = \sup_{n\in E} \mu[0,n]\,\nu[n,N] = \sup_{n\in E} \sum_{j=0}^{n} \mu_j \sum_{k=n}^{N} \frac{1}{b_k\mu_k}. \tag{3.1}$$

More precisely, we have $(4\delta)^{-1} \leqslant \lambda_0 \leqslant \delta^{-1}$. *In particular, when* $N=\infty$, *we have* $\lambda_0 = 0$ *if the process is recurrent (i.e.,* $\nu[1,\infty)=\infty$*) and* $\lambda_0>0$ *if the process is explosive (i.e., condition (1.2) does not hold).*

Proof. (a) Let $\varphi_n = \sum_{j=n}^N \nu_j =: \nu[n,N]$, $\nu_j = (b_j\mu_j)^{-1}$. To prove the lower estimate, without loss of generality, assume that $\varphi_0 < \infty$. Otherwise, $\delta=\infty$ and so the estimate is trivial. Next, let $M_n = \mu[0,n] := \sum_{k=0}^n \mu_k$. By using the summation by parts formula

$$\sum_{k=0}^{n} x_k y_k = X_n y_n - \sum_{k=0}^{n-1} X_k(y_{k+1}-y_k), \qquad X_n := \sum_{j=0}^{n} x_j, \tag{3.2}$$

in viewing the definition of δ and using the decreasing property of φ, we get

$$\begin{aligned}\sum_{j=0}^{n} \mu_j\sqrt{\varphi_j} &= M_n\sqrt{\varphi_n} + \sum_{k=0}^{n-1} M_k\big(\sqrt{\varphi_k}-\sqrt{\varphi_{k+1}}\big)\\ &\leqslant \frac{\delta}{\sqrt{\varphi_n}} + \delta\sum_{k=0}^{n-1}\frac{\sqrt{\varphi_k}-\sqrt{\varphi_{k+1}}}{\varphi_k}.\end{aligned}$$

Noting that

$$\big(\sqrt{\varphi_k}-\sqrt{\varphi_{k+1}}\big)/\varphi_k \leqslant 1/\sqrt{\varphi_{k+1}} - 1/\sqrt{\varphi_k},$$

we obtain

$$\sum_{j=0}^{n}\mu_j\sqrt{\varphi_j} \leqslant \frac{2\delta}{\sqrt{\varphi_n}}.$$

Therefore,

$$I_n\big(\sqrt{\varphi}\big) \leqslant \frac{1}{\mu_n b_n\big(\sqrt{\varphi_n}-\sqrt{\varphi_{n+1}}\big)}\cdot\frac{2\delta}{\sqrt{\varphi_n}} = \frac{2\delta}{\sqrt{\varphi_n}}\big(\sqrt{\varphi_n}+\sqrt{\varphi_{n+1}}\big) \leqslant 4\delta.$$

By part (2) of Theorem 2.4, we have $\lambda_0 \geqslant (4\delta)^{-1}$.

(b) Next, fix arbitrarily $n<m$ and let $f_i = \nu[i\vee n, m]\,\mathbb{1}_{\{i\leqslant m\}}$. Then $f\in\widetilde{\mathscr{F}}_I$. To compute $I_i(f)$, note that when $i<n$ or $i>m$, we have $f_i - f_{i+1} = 0$ but $\sum_{j\leqslant i}\mu_j f_j \geqslant \mu_0 f_0 > 0$; and when $n\leqslant i\leqslant m$, we have $f_i - f_{i+1} = \nu_i = (b_i\mu_i)^{-1}$. Hence, we have

$$I_i(f) = \begin{cases} \mu[0,n]\,\nu[n,m] + \sum_{n+1\leqslant j\leqslant i}\mu_j\,\nu[j,m], & n\leqslant i\leqslant m,\\ \infty \text{ (by convention, } 1/0=\infty), & \text{otherwise.}\end{cases}$$

Clearly, $I_i(f)$ achieves its minimum at $i = n$,

$$\inf_{i\in E} I_i(f) = \mu[0,n]\,\nu[n,m].$$

Since $n, m\,(n < m)$ are arbitrary, by letting $m \to N$ and making the supremum in n, it follows that

$$\sup_{f\in\widetilde{\mathscr{F}}_I}\inf_{i\in E} I_i(f) \geqslant \sup_{n\in E}\mu[0,n]\,\nu[n,N] = \delta.$$

By using part (2) of Theorem 2.4 again, we obtain $\lambda_0 \leqslant \delta^{-1}$. Note that in this proof, we do not preassume that $\delta < \infty$.

(c) The particular assertion for the recurrent case is obvious. The explosive case is also easy since

$$\infty > \sum_{i=0}^{\infty}\mu_i\,\nu[i,\infty) > \sum_{i=0}^{n}\mu_i\,\nu[i,\infty) > \mu[0,n]\,\nu[n,\infty)$$

for all n, and so $\delta < \infty$. $\square$

The next result is parallel to [7; Theorem 2.2], and is a typical application of parts (2) and (3) of Theorem 2.4. It provides us a way to improve step by step the estimates of λ_0. In view of Theorem 3.1, the result is meaningful only if $\delta < \infty$.

Theorem 3.2 (Approximating procedure). *Write $\nu_j = (\mu_j b_j)^{-1}$ and $\varphi_i = \nu[i,N] := \sum_{j=i}^N \nu_j$, $i \in E$.*

(1) *When $\varphi_0 < \infty$, define $f_1 = \sqrt{\varphi}$, $f_n = f_{n-1} I\!I(f_{n-1})$ and $\delta_n = \sup_{i\in E} I\!I_i(f_n)$. Otherwise, define $\delta_n \equiv \infty$. Then δ_n is decreasing in n (denote its limit by δ_∞) and*

$$\lambda_0 \geqslant \delta_\infty^{-1} \geqslant \cdots \geqslant \delta_1^{-1} \geqslant (4\delta)^{-1},$$

where δ is defined in Theorem 3.1.

(2) *For fixed $\ell, m \in E$, $\ell < m$ and $m \geqslant 1$, define*

$$f_1^{(\ell,m)} = \nu[\cdot\vee\ell, m]\,\mathbb{1}_{\leqslant m},$$
$$f_n^{(\ell,m)} = \mathbb{1}_{\leqslant m}\, f_{n-1}^{(\ell,m)}\, I\!I\big(f_{n-1}^{(\ell,m)}\big), \quad n \geqslant 2,$$

where $\mathbb{1}_{\leqslant m}$ is the indicator of the set $\{0,1,\ldots,m\}$, and then define

$$\delta_n' = \sup_{\ell,m:\,\ell<m}\min_{i\leqslant m} I\!I_i\big(f_n^{(\ell,m)}\big).$$

Then δ_n' is increasing in n (denote its limit by δ_∞') and

$$\delta^{-1} \geqslant {\delta_1'}^{-1} \geqslant \cdots \geqslant {\delta_\infty'}^{-1} \geqslant \lambda_0.$$

Next, define

$$\bar{\delta}_n = \sup_{\ell,m:\,\ell<m}\frac{\big\|f_n^{(\ell,m)}\big\|^2}{D\big(f_n^{(\ell,m)}\big)}, \qquad n \geqslant 1.$$

Then $\bar{\delta}_n^{-1} \geqslant \lambda_0$, $\bar{\delta}_{n+1} \geqslant \delta_n'$ for all $n \geqslant 1$ and $\bar{\delta}_1 = \delta_1'$.

As the first step of the above approximation, we obtain the following improvement of Theorem 3.1.

Corollary 3.3 (Improved estimates). *We have*

$$\delta^{-1} \geqslant {\delta_1'}^{-1} \geqslant \lambda_0 \geqslant \delta_1^{-1} \geqslant (4\delta)^{-1}, \tag{3.3}$$

where

$$\begin{aligned}\delta_1 &= \sup_{i\in E} \frac{1}{\sqrt{\varphi_i}} \sum_{k\in E} \mu_k \varphi_{i\vee k}\sqrt{\varphi_k}\\ &= \sup_{i\in E}\bigg[\sqrt{\varphi_i}\sum_{k=0}^{i}\mu_k\sqrt{\varphi_k} + \frac{1}{\sqrt{\varphi_i}}\sum_{i+1\leqslant k<N+1}\mu_k\varphi_k^{3/2}\bigg].\end{aligned} \tag{3.4}$$

$$\delta_1' = \sup_{\ell\in E}\frac{1}{\varphi_\ell}\sum_{k\in E}\mu_k\varphi_{k\vee\ell}^2 = \sup_{\ell\in E}\bigg[\varphi_\ell\,\mu[0,\ell] + \frac{1}{\varphi_\ell}\sum_{k=\ell+1}^{N}\mu_k\varphi_k^2\bigg] \in [\delta,\, 2\delta]. \tag{3.5}$$

Proofs of Theorem 3.2 *and Corollary* 3.3. (a) First, we prove part (1) of Theorem 3.2. Noting that if $\varphi_0 = \infty$, then $\delta = \infty$ and $\delta_n = \infty$ for all $n \geqslant 1$, the assertion becomes trivial in view of Theorem 3.1. Thus, we can assume that $\varphi_0 < \infty$.

By (2.23), we have

$$\delta_1 = \sup_{i\in E} \mathit{II}_i(f_1) \leqslant \sup_{i\in E} I_i(f_1).$$

Proof (a) of Theorem 3.1 shows that the last one is bounded from above by 4δ. This gives us the lower bound of δ_1^{-1} as required.

We now prove the monotonicity of $\{\delta_n\}$. By induction, assume that $f_n < \infty$ and $\delta_n < \infty$. Then $f_{n+1} < \infty$. Note that

$$\begin{aligned}\sum_{j\leqslant k}\mu_j f_{n+1}(j) &= \sum_{j\leqslant k}\mu_j f_n(j) f_{n+1}(j)/f_n(j)\\ &\leqslant \sup_{i\in E}\mathit{II}_i(f_n)\sum_{j\leqslant k}\mu_j f_n(j)\\ &= \delta_n\sum_{j\leqslant k}\mu_j f_n(j).\end{aligned}$$

Multiplying both sides by ν_k and making a summation of k from i to N, by (2.14), it follows that

$$f_{n+2}(i) \leqslant \delta_n f_{n+1}(i).$$

Because $\delta_n < \infty$ and $f_{n+1}(i) < \infty$, we obtain $f_{n+2} < \infty$ and $\mathit{II}_i(f_{n+1}) \leqslant \delta_n < \infty$. Now, making the supremum over i, we obtain $\delta_{n+1} \leqslant \delta_n < \infty$.

We have thus proved part (1) of Theorem 3.2.

(b) To prove the monotonicity of δ_n' given in part (2) of Theorem 3.2, we use the proportional property twice:

$$\begin{aligned}\min_{i\leqslant m}\big[f_{n+1}^{(\ell,m)}/f_n^{(\ell,m)}\big](i) &= \min_{i\leqslant m}\sum_{j=i}^{m}\nu_j\sum_{k\leqslant j}\mu_k f_n^{(\ell,m)}(k)\bigg/\sum_{j=i}^{m}\nu_j\sum_{k\leqslant j}\mu_k f_{n-1}^{(\ell,m)}(k)\\ &\geqslant \min_{i\leqslant m}\sum_{k\leqslant i}\mu_k f_n^{(\ell,m)}(k)\bigg/\sum_{k\leqslant i}\mu_k f_{n-1}^{(\ell,m)}(k)\\ &\geqslant \min_{i\leqslant m} f_n^{(\ell,m)}(i)\big/f_{n-1}^{(\ell,m)}(i).\end{aligned}$$

This implies that $\delta'_{n+1} \geqslant \delta'_n$.

By part (2) of Theorem 2.4, we also have $\delta'_n \leqslant \lambda_0^{-1}$ for all $n \geqslant 1$. The assertion that $\bar{\delta}_n \leqslant \lambda_0^{-1}$ is obvious. Next, let $f = f_n^{(\ell,m)}$. Then $g := \mathbb{1}_{\operatorname{supp}(f)} f I\!I(f) = f_{n+1}^{(\ell,m)}$. As a consequence of (2.30), we obtain $\bar{\delta}_{n+1} \geqslant \delta'_n$.

We have thus proved part (2) of Theorem 3.2 except the last assertion that $\bar{\delta}_1 = \delta'_1$.

(c) We now prove (3.4) and $\delta'_1 \geqslant \delta$. By (2.15), we have

$$\begin{aligned}
f_{n+1}(i) &= \sum_{k\in E} \mu_k f_n(k)\,\nu[i\vee k, N]\\
&= \sum_{k\in E} \mu_k f_n(k)\,\varphi_{i\vee k}\\
&= \varphi_i \sum_{k=0}^{i} \mu_k f_n(k) + \sum_{i+1\leqslant k<N+1} \mu_k\,\varphi_k f_n(k).
\end{aligned}\tag{3.6}$$

In particular, with $f_1 = \sqrt{\varphi}$, we get

$$f_2(i) = \varphi_i \sum_{k=0}^{i} \mu_k \sqrt{\varphi_k} + \sum_{i+1\leqslant k<N+1} \mu_k \varphi_k^{3/2}.\tag{3.7}$$

From this, we obtain (3.4).

To prove $\delta'_1 \geqslant \delta$, we need some preparation. As an analog of (3.6), we have

$$f_{n+1}^{(\ell,m)}(i) = \mathbb{1}_{\{i\leqslant m\}} \sum_{k\leqslant m} \mu_k f_n^{(\ell,m)}(k)\,\nu[i\vee k, m].\tag{3.8}$$

In particular,

$$f_2^{(\ell,m)}(i) = \mathbb{1}_{\{i\leqslant m\}} \sum_{k\leqslant m} \mu_k\,\nu[k\vee \ell, m]\,\nu[i\vee k, m].\tag{3.9}$$

Since the right-hand side is decreasing in i for $i \leqslant \ell < m$, $f_1^{(\ell,m)}(i) = f_1^{(\ell,m)}(\ell)$ for all $i \leqslant \ell$, and $f_1^{(\ell,m)}(i) = 0$ for $i > m$, it follows that

$$\begin{aligned}
\min_{i\leqslant m} I\!I_i\big(f_1^{(\ell,m)}\big) &= \min_{\ell\leqslant i\leqslant m} I\!I_i\big(f_1^{(\ell,m)}\big)\\
&= \min_{\ell\leqslant i\leqslant m} \frac{1}{\nu[i,m]} \sum_{j=i}^{m} \nu_j \sum_{k\leqslant j} \mu_k\,\nu[k\vee\ell, m]\,\mathbb{1}_{\{k\leqslant m\}}\\
&= \min_{\ell\leqslant i\leqslant m} \sum_{j=i}^{m} \nu_j \sum_{k\leqslant j} \mu_k\,\nu[k\vee\ell, m] \bigg/ \sum_{j=i}^{m} \nu_j\\
&\geqslant \min_{\ell\leqslant i\leqslant m} \sum_{k\leqslant i} \mu_k\,\nu[k\vee\ell, m] \quad \Big[= \inf_{i\in E} I_i\big(f_1^{(\ell,m)}\big)\Big].
\end{aligned}$$

Here in the last step, we have used the proportional property. Since the sum on the right-hand side is increasing in i, it is clear that

$$\min_{\ell\leqslant i\leqslant m}\sum_{k\leqslant i}\mu_k\,\nu[k\vee\ell,m]=\nu[\ell,m]\sum_{k\leqslant\ell}\mu_k=\mu[0,\ell]\,\nu[\ell,m].$$

We have thus proved that

$$\delta_1'=\sup_{\ell<m}\min_{i\leqslant m}I\!I_i\big(f_1^{(\ell,m)}\big)\geqslant\sup_{\ell<m}\mu[0,\ell]\,\nu[\ell,m]\geqslant\sup_{\ell\in E}\mu[0,\ell]\,\varphi_\ell=\delta.$$

A different proof of this is given in proof (d) below.

(d) We now compute δ_1'. Note that by (3.9), we have

$$f_2^{(\ell,m)}(i)=\mathbb{1}_{\{i\leqslant m\}}\sum_{k=0}^m\mu_k\,\nu[k\vee\ell,m]\,\nu[i\vee k,m].$$

Since $f_2^{(\ell,m)}(i)$ is decreasing in i and $f_1^{(\ell,m)}(i)$ is a constant on $\{0,1,\ldots,\ell\}$, it is clear that $\min_{i\leqslant m}f_2^{(\ell,m)}(i)/f_1^{(\ell,m)}(i)=\min_{\ell\leqslant i\leqslant m}f_2^{(\ell,m)}(i)/f_1^{(\ell,m)}(i)$. Besides, when $\ell\leqslant i\leqslant m$, we have

$$f_2^{(\ell,m)}(i)=\nu[\ell,m]\nu[i,m]\sum_{k=0}^{\ell}\mu_k+\nu[i,m]\sum_{\ell+1\leqslant k\leqslant i}\mu_k\nu[k,m]+\sum_{i+1\leqslant k\leqslant m}\mu_k\nu[k,m]^2.$$

It follows that

$$\min_{i\leqslant m}\frac{f_2^{(\ell,m)}(i)}{f_1^{(\ell,m)}(i)}=\min_{\ell\leqslant i\leqslant m}\bigg[\nu[\ell,m]\sum_{k=0}^{\ell}\mu_k+\sum_{k=\ell+1}^{i}\mu_k\nu[k,m]+\frac{1}{\nu[i,m]}\sum_{k=i+1}^{m}\mu_k\nu[k,m]^2\bigg].$$

We show that the sum on the right-hand side is increasing in i. That is,

$$\begin{aligned}&\sum_{\ell+1\leqslant k\leqslant i}\mu_k\nu[k,m]+\frac{1}{\nu[i,m]}\sum_{k=i+1}^{m}\mu_k\nu[k,m]^2\\&\qquad\leqslant\sum_{k=\ell+1}^{i+1}\mu_k\nu[k,m]+\frac{1}{\nu[i+1,m]}\sum_{i+2\leqslant k\leqslant m}\mu_k\nu[k,m]^2,\qquad \ell\leqslant i\leqslant m-1.\end{aligned}$$

Collecting the terms, this is equivalent to

$$\frac{1}{\nu[i,m]}\mu_{i+1}\nu[i+1,m]^2\leqslant\mu_{i+1}\nu[i+1,m]+\bigg(\frac{1}{\nu[i+1,m]}-\frac{1}{\nu[i,m]}\bigg)\sum_{k=i+2}^{m}\mu_k\nu[k,m]^2.$$

Now, the conclusion becomes obvious because by the decreasing property of $\nu[i,m]$ in i, the first term is controlled by the second, and the last one is nonnegative. We have thus obtained that

$$\min_{\ell\leqslant i\leqslant m}\frac{f_2^{(\ell,m)}(i)}{f_1^{(\ell,m)}(i)}=\nu[\ell,m]\sum_{k=0}^{\ell}\mu_k+\frac{1}{\nu[\ell,m]}\sum_{k=\ell+1}^{m}\mu_k\nu[k,m]^2.\tag{3.10}$$

As will be seen soon that the right-hand side is increasing in $m\,(>\ell)$, hence, we obtain

$$\delta_1' = \sup_{\ell<m} \min_{\ell\leqslant i\leqslant m} \frac{f_2^{(\ell,m)}(i)}{f_1^{(\ell,m)}(i)} = \sup_{\ell\in E}\bigg[\varphi_\ell \sum_{k=0}^{\ell}\mu_k + \frac{1}{\varphi_\ell}\sum_{\ell+1\leqslant k<N+1}\mu_k\varphi_k^2\bigg]. \tag{3.11}$$

From this, it follows once again that $\delta_1'\geqslant\delta$. We now turn to prove the monotone property:

$$\begin{aligned}\mu[0,\ell]\,\nu[\ell,m+1] &+ \frac{1}{\nu[\ell,m+1]}\sum_{i=\ell+1}^{m+1}\mu_i\,\nu[i,m+1]^2\\ &\geqslant \mu[0,\ell]\,\nu[\ell,m] + \frac{1}{\nu[\ell,m]}\sum_{i=\ell+1}^{m}\mu_i\,\nu[i,m]^2.\end{aligned}$$

Equivalently,

$$\mu[0,\ell]\,\nu_{m+1} + \frac{\mu_{m+1}}{\nu[\ell,m+1]}\,\nu_{m+1}^2 + \sum_{i=\ell+1}^{m}\mu_i\bigg(\frac{\nu[i,m+1]^2}{\nu[\ell,m+1]} - \frac{\nu[i,m]^2}{\nu[\ell,m]}\bigg)\geqslant 0.$$

This becomes obvious since the term in the last bracket is positive:

$$\frac{\nu[i,m+1]^2}{\nu[i,m]^2} = \bigg(1+\frac{\nu_{m+1}}{\nu[i,m]}\bigg)^2 > 1+\frac{\nu_{m+1}}{\nu[\ell,m]} = \frac{\nu[\ell,m+1]}{\nu[\ell,m]}, \qquad \ell\leqslant i\leqslant m.$$

(e) To show that $\delta_1'\leqslant 2\delta$, assume $\delta<\infty$. By using the summation by parts formula (3.2) with $x_k=\mu_k$, $X_k=\sum_{j=0}^k\mu_j$, and $y_k=\varphi_{k\vee i}^2$, we get

$$\begin{aligned}\sum_{k=0}^{M}\mu_k\varphi_{k\vee i}^2 &= \varphi_M^2 X_M + \sum_{k=0}^{M-1}X_k\big[\varphi_{k\vee i}^2-\varphi_{(k+1)\vee i}^2\big]\\ &= \varphi_M^2 X_M + \sum_{k=i}^{M-1}X_k\big[\varphi_k^2-\varphi_{k+1}^2\big]\\ &= \varphi_M^2 X_M + \sum_{k=i}^{M-1}X_k\nu_k(\varphi_k+\varphi_{k+1})\\ &< \varphi_M^2 X_M + 2\sum_{k=i}^{M-1}X_k\nu_k\varphi_k\\ &\leqslant \delta\varphi_M + 2\delta\sum_{k=i}^{M-1}\nu_k \quad (\text{since } X_k\varphi_k\leqslant\delta), \qquad i<M<N+1.\end{aligned}$$

If $N=\infty$, letting $M\to N$, it follows that

$$\sum_{k=0}^{N}\mu_k\varphi_{k\vee i}^2\leqslant 2\delta\varphi_i.$$

The same conclusion holds in the case that $N<\infty$ since

$$\delta\varphi_N+2\delta\sum_{k=i}^{N-1}\nu_k<2\delta\sum_{k=i}^{N}\nu_k=2\delta\varphi_i.$$

Hence,

$$\delta_1'=\sup_{i\in E}\frac{1}{\varphi_i}\sum_{k=0}^{N}\mu_k\varphi_{k\vee i}^2\leqslant 2\delta.$$

(f) Now, it remains to compute $\bar{\delta}_1$. Since $f_1^{(\ell,m)}(i)=\nu[i\vee\ell,m]\,\mathbb{1}_{\{i\leqslant m\}}$, we have

$$\big\|f_1^{(\ell,m)}\big\|^2=\sum_i\mu_i\,\nu[i\vee\ell,m]^2\mathbb{1}_{\{i\leqslant m\}}=\mu[0,\ell]\,\nu[\ell,m]^2+\sum_{i=\ell+1}^{m}\mu_i\,\nu[i,m]^2,$$

and

$$\begin{aligned}D\big(f_1^{(\ell,m)}\big)&=\sum_i\mu_ib_i\big(f_1^{(\ell,m)}(i+1)-f_1^{(\ell,m)}(i)\big)^2\\&=\sum_{i=\ell}^{m-1}\mu_ib_i\big(\nu[i+1,m]-\nu[i,m]\big)^2+\mu_mb_m\nu_m^2\\&=\sum_{i=\ell}^{m-1}\nu_i+\nu_m\\&=\nu[\ell,m].\end{aligned}$$

Thus,

$$\frac{\big\|f_1^{(\ell,m)}\big\|^2}{D\big(f_1^{(\ell,m)}\big)}=\mu[0,\ell]\,\nu[\ell,m]+\frac{1}{\nu[\ell,m]}\sum_{i=\ell+1}^{m}\mu_i\,\nu[i,m]^2.$$

Hence, we have returned to (3.10). Since the right-hand side is increasing in m as we have seen in the proof of (3.11), we obtain

$$\bar{\delta}_1=\sup_{\ell<m}\frac{\big\|f_1^{(\ell,m)}\big\|^2}{D\big(f_1^{(\ell,m)}\big)}=\delta_1'.\quad\square$$

To conclude this section, we present some examples to illustrate the power of our results. The first one is standard having constant rates.

Example 3.4. *Let $b_i\equiv b>0\,(i\geqslant 0)$, $a_i\equiv a>0\,(i\geqslant 1)$, $b>a$. Then*

(1) $\lambda_0=\big(\sqrt{a}-\sqrt{b}\big)^2$ *with eigenfunction g:*

$$g_n=\Big(\frac{a}{b}\Big)^{n/2}\bigg(n+1-n\sqrt{\frac{a}{b}}\bigg),\qquad n\geqslant 0,\qquad g\notin L^1(\mu)\cup L^2(\mu).$$

(2) $\delta=b(b-a)^{-2}$, $\delta_1'=(a+b)(b-a)^{-2}=\bar{\delta}_1>\delta$, *and* $\delta_1=\lambda_0^{-1}$ *which is exact. Note that* $\delta_1/\delta_1'<2$ *whenever* $a\neq b$ *and* $\lim_{b\to a}\delta_1/\delta_1'=2$. *When* $a=b$, *we have* $\lambda_0=\delta_1^{-1}={\delta_1'}^{-1}=0$.

Proof. (a) First, we have $\mu_n = (b/a)^n$, $n \geqslant 0$. Hence,

$$\sum_n \mu_n = \infty, \qquad \sum_n \mu_n g_n \geqslant \sum_n \mu_n g_n^2 \geqslant \sum_n 1 = \infty.$$

Next, since

$$\nu_i = \frac{1}{\mu_i b_i} = \frac{1}{b}\left(\frac{a}{b}\right)^i,$$

we have

$$\varphi_\ell = \sum_{i \geqslant \ell} \nu_i = \frac{1}{b-a}\left(\frac{a}{b}\right)^\ell$$

and then

$$\sum_{i=0}^\infty \mu_i \sum_{k=i}^\infty \frac{1}{b_k \mu_k} = \sum_{i=0}^\infty \mu_i \varphi_i = \infty.$$

Hence, (1.2) holds. It is easy to check that (2.12) holds:

$$\sum_{n=0}^\infty \mu_n g_n \nu[n,\infty) = \frac{1}{b-a}\sum_{n=0}^\infty \left(\frac{a}{b}\right)^{n/2}\left(n+1-n\sqrt{\frac{a}{b}}\right) = \frac{1}{\lambda_0}.$$

(b) To study λ_0, according to (a), the Dirichlet form is regular and so the condition “$f \in \mathscr{K}$” in the definition of λ_0 can be ignored. Thus,

$$\lambda_0 = \inf_{\|f\|=1} D(f) = b \inf_{\|f\|=1} \sum_{i \geqslant 0} \mu_i (f_{i+1} - f_i)^2.$$

It suffices to consider the case that $b = 1$. Write $\gamma = b/a > 1$. Then we have

$$g_k = \gamma^{-k/2}(k+1-k\gamma^{-1/2}), \quad \mu_k = \gamma^k, \quad \nu_k = \gamma^{-k}, \quad \varphi_k = \gamma^{-k+1}/(\gamma-1),$$

and the required quantities are reduced to

$$\lambda_0 = \frac{(\sqrt{\gamma}-1)^2}{\gamma}, \quad \delta = \frac{\gamma^2}{(\gamma-1)^2}, \quad \delta_1 = \frac{\gamma}{(\sqrt{\gamma}-1)^2}, \quad \delta_1' = \frac{\gamma(\gamma+1)}{(\gamma-1)^2}.$$

Now, to prove part (1) of Example 3.4, write $\xi = (\sqrt{\gamma}-1)^2\gamma^{-1}$ for distinguishing with λ_0. Since (g, ξ) satisfies the eigenequation, applying anyone of the variational formulas for the lower estimate given in Theorem 2.4 with $f_i = g_i$ or

$$v_i = \frac{g_{i+1}}{g_i} = \sqrt{\frac{a}{b}}\left(1 + \frac{1-\sqrt{a/b}}{1+i(1-\sqrt{a/b}\,)}\right) = \gamma^{-1/2}\left(1 + \frac{1-\gamma^{-1/2}}{1+i(1-\gamma^{-1/2})}\right),$$

it follows that $\lambda_0 \geqslant \xi$. We have seen that the equality sign holds once $g \in L^2(\mu)$. Unfortunately, we are now out of this case. Therefore, we need to show that

$\lambda_0 \leqslant \xi$. To do so, one may use the truncated function of g: $g_i^{(m)} = g_i \mathbb{1}_{\{i \leqslant m\}}$. Then by the Stolz theorem, we have

$$\lambda_0 \leqslant \lim_{m\to\infty} \frac{D(g^{(m)})}{\|g^{(m)}\|^2} = \lim_{m\to\infty} \left[b_m + \frac{\mu_{m-1}b_{m-1}}{\mu_m}\left(1 - \frac{2g_{m-1}}{g_m}\right)\right]. \tag{3.12}$$

The last limit equals ξ. Alternatively, noting that the leading order of $g \notin L^2(\mu)$ is $\gamma^{-k/2}$, one may adopt the test function $f_i = z^{-i/2}$ for $z > \gamma$. Then $f \in L^2(\mu)$. The required assertion follows by computing $D(f)/\|f\|^2$ and then letting $z \downarrow \gamma$. This proof benefits very much from the explicitly known expression of λ_0.

(c) The computation of δ is easy:

$$\delta = \sup_{n\geqslant 0} \varphi_n \sum_{j=0}^{n} \mu_j = \frac{1}{(\gamma-1)^2} \sup_{n\geqslant 0} \gamma^{-n+1}\big(\gamma^{n+1} - 1\big) = \frac{\gamma^2}{(\gamma-1)^2}.$$

(d) To compute δ_1, by (3.7), we have

$$\begin{aligned}
f_2(i) &= \varphi_i \sum_{k=0}^{i} \mu_k \sqrt{\varphi_k} + \sum_{k=i+1}^{\infty} \mu_k \varphi_k^{3/2} \\
&= \frac{1}{(\gamma-1)^{3/2}} \left\{\gamma^{-i+1} \sum_{k=0}^{i} \gamma^{k/2+1/2} + \sum_{k\geqslant i+1} \gamma^{-k/2+3/2}\right\} \\
&= \frac{\gamma^{-i/2+3/2}}{(\gamma-1)^{3/2}(\sqrt{\gamma}-1)} \big(\sqrt{\gamma} - \gamma^{-i/2} + 1\big).
\end{aligned}$$

Therefore, we obtain

$$\delta_1 = \sup_{i\geqslant 0} \frac{f_2(i)}{f_1(i)} = \frac{\gamma}{(\gamma-1)(\sqrt{\gamma}-1)} \big(\sqrt{\gamma}+1\big) = \frac{\gamma}{(\sqrt{\gamma}-1)^2} = \frac{1}{\lambda_0}.$$

Noting that even if neither f_1 nor f_2 is the eigenfunction, we still obtain the sharp estimate.

(e) To compute δ_1', by (3.5), we have

$$\begin{aligned}
\delta_1' &= \sup_{\ell\in E} \left[\varphi_\ell \sum_{k=0}^{\ell} \mu_k + \frac{1}{\varphi_\ell} \sum_{k\geqslant \ell+1} \mu_k \varphi_k^2\right] \\
&= \frac{1}{\gamma-1} \sup_{\ell\in E} \left[\gamma^{-\ell+1} \sum_{k=0}^{\ell} \gamma^k + \gamma^{\ell+1} \sum_{k\geqslant \ell+1} \gamma^{-k}\right] \\
&= \frac{1}{(\gamma-1)^2} \sup_{\ell\in E} \left[\gamma^2 - \gamma^{-\ell+1} + \gamma\right] \\
&= \frac{\gamma(\gamma+1)}{(\gamma-1)^2}. \quad \square
\end{aligned}$$

The next example is a typical linear model for which, interestingly, we have a very simple and common eigenfunction. Moreover, the eigenvalue λ_0 is determined by the constant term 2γ in the rates, but not the difference of the coefficients of the leading term i, as in the ergodic case (cf. Example 6.8 below).

Example 3.5. *Let* $b_i = 2(i+\gamma)\,(i \geqslant 0)$, $\gamma > 0$, $a_i = i\,(i \geqslant 1)$. *Then*

(1) $\lambda_0 = \gamma$, $g_n = 2^{-n}$ *for* $n \geqslant 0$, *and* $g \in L^2(\mu)$.

(2) *When* $\gamma = 1$, *we have* $\delta = \log 2 \approx 0.69$, $\delta_1' \approx 0.84$, *and* $\delta_1 \approx 1.09$. *Then* $\delta_1/\delta_1' \approx 1.3 < 2$.

Proof. The uniqueness condition (1.2) is trivial since the birth rates are linear:

$$\sum_{k=0}^{\infty} \frac{1}{b_k \mu_k} \sum_{i=0}^{k} \mu_i = \sum_{k=0}^{\infty} \left[\frac{1}{b_k} + \frac{1}{b_k \mu_k} \sum_{i=0}^{k-1} \mu_i \right] \geqslant \sum_{k=0}^{\infty} \frac{1}{b_k} = \infty.$$

(a) Because

$$\mu_0 = 1, \qquad \mu_n = \frac{2^n \gamma(1+\gamma)\cdots(n-1+\gamma)}{n!}, \qquad n \geqslant 1,$$

it follows that $\mu_n > \gamma 2^n/n$ and so $\sum_n \mu_n = \infty$. Next, since

$$\mu_n b_n = \frac{2^{n+1}\gamma(1+\gamma)\cdots(n+\gamma)}{n!} > \gamma 2^{n+1},$$

we have $\sum_n (\mu_n b_n)^{-1} < \infty$. Furthermore, we have

$$\sum_{n=0}^{\infty} \mu_n g_n^2 = \sum_{n=0}^{\infty} \frac{2^{-n}\gamma(1+\gamma)\cdots(n-1+\gamma)}{n!}.$$

The ratio test tells us $g \in L^2(\mu)$. Since λ_0 is explicit and $g \in L^2(\mu)$, it is simple to check that (g_n) is the eigenfunction of λ_0. Hence, the proof of part (1) is done. For this example, the sequence (v_i) takes a simple form: $v_i \equiv 1/2$.

(b) When $\gamma = 1$, we have $\lambda_0 = 1$,

$$\mu_i = 2^i, \qquad \mu_i b_i = (i+1)2^{i+1}, \qquad \varphi_i = \sum_{k \geqslant i+1} \frac{1}{2^i i}, \qquad i \geqslant 0.$$

In particular, $\varphi_0 = \log 2$, $\varphi_1 = \log 2 - 1/2$. Numerical computations show that the supremum in the definition of δ, δ_1' and δ_1 are attained at 0, 0 and 1, respectively, and moreover,

$$\begin{aligned}
\delta &= \varphi_0 \mu_0 = \varphi_0 = \log 2 \approx 0.69,\\
\delta_1' &= \varphi_0 \mu_0 + \frac{1}{\varphi_0} \sum_{k \geqslant 1} \mu_k \varphi_k^2 = \log 2 + \frac{1}{\log 2} \sum_{k \geqslant 1} 2^k \varphi_k^2 \approx 0.84,\\
\delta_1 &= \sqrt{\varphi_1}\,(\mu_0 \sqrt{\varphi_0} + \mu_1 \sqrt{\varphi_1}\,) + \frac{1}{\sqrt{\varphi_1}} \sum_{k \geqslant 2} \mu_k \varphi_k^{3/2}\\
&= 2\log 2 - 1 + \frac{1}{2}\sqrt{(2\log 2)(2\log 2 - 1)} + \sqrt{\frac{2}{2\log 2 - 1}} \sum_{k \geqslant 2} 2^k \varphi_k^{3/2}\\
&\approx 1.09.
\end{aligned}$$

We have thus proved part (2) of the conclusion. $\square$

The next example is often used in the study of convergence rates. For which, the first eigenfunction is unknown but λ_0 can still be computed.

Example 3.6. *Let $b_i = (i+1)^2$ and $a_i = i^2$. Then $\delta = \pi^2/6 \approx 1.64$, $\delta_1' \approx 2.19$, and $\delta_1 = 4$ which is sharp ($\lambda_0 = 1/4$). Besides, $\delta_1/\delta_1' \approx 1.83 < 2$.*

Proof. (a) Since $\mu_i \equiv 1$, $\nu_i = (i+1)^{-2}$, we have $\mu[0,i] = i+1$ and $\varphi_i = \sum_{j\geqslant i+1} j^{-2}$. For δ and δ_1', the supremum is attained at 0, therefore,

$$\delta = \varphi_0 = \sum_{k\geqslant 1} \frac{1}{k^2} = \frac{\pi^2}{6},$$

and

$$\delta_1' = \frac{1}{\varphi_0} \sum_{k=0}^{\infty} \varphi_k^2 \approx 2.19.$$

(b) For δ_1, the supremum is attained at ∞ and is equal to 4. By Corollary 3.3, this means that $\lambda_0 \geqslant 1/4$. This can be also deduced by part (1) of Theorem 2.4 with $v_i = 1-(2i+4)^{-1}$ for which the minimum of $R_i(v)$ is attained at $i=0$ and $i=\infty$. It is even more simpler to use $v_i = 1-(2i+3)^{-1}$. Next, it is known that $\lambda_0 \leqslant 1/4$ (cf. Example 5.5 below), hence, the estimate is sharp. A direct proof for the upper estimate goes as follows. Since the lower estimate is sharp, it indicates to use the test function

$$f_i = \bigg(\sum_{j=i}^{\infty} \frac{1}{(j+1)^2}\bigg)^{1/2} \sim \frac{1}{\sqrt{i+1}}.$$

However, the last function is not in $L^2(\mu)$, and so one needs an approximating procedure. Now, a carefully designed test function is the following:

$$f_i^{(\alpha)} = \frac{1}{\sqrt{(i+1)\alpha^{i+1}}}, \qquad \alpha > 1.$$

Then

$$\begin{aligned}
\mu\big(f^{(\alpha)\,2}\big) &= \sum_{i=0}^{\infty} \frac{1}{(i+1)\alpha^{i+1}} = \sum_{i=1}^{\infty} \frac{1}{i\alpha^i} = \log[\alpha(\alpha-1)^{-1}] < \infty,\\
D\big(f^{(\alpha)}\big) &= \sum_{i=0}^{\infty} (i+1)^2 \bigg[\frac{1}{\sqrt{(i+2)\alpha^{i+2}}} - \frac{1}{\sqrt{(i+1)\alpha^{i+1}}}\bigg]^2\\
&= \sum_{i=1}^{\infty} \frac{i^2}{\alpha^i} \bigg[\frac{1}{\sqrt{(i+1)\alpha}} - \frac{1}{\sqrt{i}}\bigg]^2\\
&= \sum_{i=1}^{\infty} \frac{i}{(i+1)\alpha^{i+1}} \frac{[(i+1)\alpha - i]^2}{[\sqrt{(i+1)\alpha} + \sqrt{i}\,]^2}\\
&\leqslant \frac{1}{4} \sum_{i=1}^{\infty} \frac{1}{(i+1)\alpha^{i+1}} [(i+1)\alpha - i]^2\\
&= \frac{1}{4}(2 + \log[\alpha(\alpha-1)^{-1}]).
\end{aligned}$$

The required assertion now follows from

$$\lambda_0 \leqslant \frac{2 + \log[\alpha(\alpha-1)^{-1}]}{4\log[\alpha(\alpha-1)^{-1}]} \to \frac{1}{4} \quad \text{as } \alpha \downarrow 1. \quad \square$$

The last example below does not satisfy the non-explosive condition (1.2).

Example 3.7. *Let $b_i = (i+1)^4$ and $a_i = i(i-1/2)(i^2+3i+3)$. Then $\sum_i \mu_i < \infty$, $\sum_i \nu_i < \infty$, $\lambda_0 = 1/2$, $\delta \approx 1.83$, $\delta_1' \approx 1.9$, and $\delta_1 \approx 2$. Moreover, $\delta_1/\delta_1' \approx 1.05 < 2$.*

Proof. A simple computation shows that

$$\mu_i = \frac{i!^3}{\prod_{k=1}^i (k-1/2)(k^2+3k+3)}, \qquad \nu_i = \frac{\prod_{k=1}^i (k-1/2)(k^2+3k+3)}{(i+1)(i+1)!^3}.$$

From this, it follows that $\sum_i \mu_i < \infty$ and $\sum_i \nu_i < \infty$, as an application of the typical Kummer's test: for a positive sequence $\{x_n\}$, $\sum_n x_n$ converges or diverges according to $\kappa > 1$ or $\kappa < 1$, respectively, where

$$\kappa = \lim_{n\to\infty} n\left(\frac{x_n}{x_{n+1}} - 1\right). \tag{3.13}$$

For each of δ, δ_1' and δ_1, the supremum is attained at 0.

To see that $\lambda_0 = 1/2$, first we check that $R_i(v) \equiv 1/2$ for

$$v_i = 1 - \frac{1}{2(i+1)}.$$

This gives us $\lambda_0 \geqslant 1/2$ by part (1) of Theorem 2.4. Since the corresponding eigenfunction g,

$$g_i = \prod_{k=0}^{i-1} v_k = \frac{(2i-1)!}{2^{2i-1} i (i-1)!^2}, \quad i \geqslant 1, \qquad g_0 = 1,$$

decreases strictly to 0 and $\sum_i \mu_i < \infty$, we have $g \in L^2(\mu)$. Now, because $-\Omega g = \lambda_0 g$, $g_\infty = 0$, and $D(f) = -(g, \Omega g)$, it follows that $\lambda_0 = 1/2$ by (2.18). $\square$

4. Absorbing (Dirichlet) boundary at origin and reflecting (Neumann) boundary at infinity

This section deals with the second case of the boundary conditions. The process has state space $E = \{i : 1 \leqslant i < N+1\}$ ($N \leqslant \infty$), birth rates $b_i > 0$ but $b_N = 0$ if $N < \infty$, and death rates $a_i > 0$. The rate $a_1 > 0$ is regarded as a killing from 1. Define

$$\lambda_0 = \inf\{D(f)/\mu(f^2) : f \neq 0,\ D(f) < \infty\}, \tag{4.1}$$

where $\mu(f) = \sum_{k\in E} \mu_k f_k$, and

$$D(f) = \sum_{k\in E} \mu_k a_k (f_k - f_{k-1})^2, \qquad f_0 := 0,$$
$$\mu_1 = 1, \qquad \mu_k = \frac{b_1 \cdots b_{k-1}}{a_2 \cdots a_k}, \quad 2 \leqslant k < N+1.$$

The constant $\lambda_0^{(4.1)}$ describes the optimal constant $C = \lambda_0^{-1}$ in the following *weighted Hardy inequality*:

$$\mu\big(f^2\big) \leqslant C D(f), \qquad f_0 = 0$$

(cf. [9]). In other words, we are studying the discrete version of the weighted Hardy inequality in this section. To save the notation, in this and the subsequent sections, we use the same notation λ_0, I, $I\!I$, R and so on as in Section 2. Each of them plays a similar role but may have different meaning in different sections.

To study λ_0, as in Section 2, we need some parallel notation originally introduced in [3, 7]:

$$I_i(f) = \frac{1}{\mu_i a_i (f_i - f_{i-1})} \sum_{j=i}^{N} \mu_j f_j, \qquad I\!I_i(f) = \frac{1}{f_i} \sum_{j=1}^{i} \frac{1}{\mu_j a_j} \sum_{k=j}^{N} \mu_k f_k.$$

Here, for the first operator, we adopt the convention: $f_0 = 0$. The second one can be re-written as

$$I\!I_i(f) = \frac{1}{f_i} \sum_{k=1}^{N} \mu_k f_k \, \nu[1, i\wedge k], \qquad \nu[\ell, m] = \sum_{j=\ell}^{m} \nu_j, \qquad \nu_j = \frac{1}{\mu_j a_j}.$$

Next, define

$$R_i(v) = a_i\big(1 - v_{i-1}^{-1}\big) + b_i(1 - v_i), \qquad i \in E,\ v_0 := \infty$$

(v_N is free if $N < \infty$ since $b_N = 0$) and

$$\mathscr{F}_{I\!I} = \{f\colon f_i > 0 \text{ for all } i \in E\},$$
$$\mathscr{F}_{I} = \big\{f : f > 0 \text{ and is strictly increasing on } E\big\},$$
$$\mathscr{V}_1 = \{v : v_i > 1 \text{ for all } i \in E\}.$$

The modifications of $\mathscr{F}_{I\!I}$ and $\mathscr{F}_I$ are as follows:

$$\widetilde{\mathscr{F}}_{I\!I} = \{f : \text{ there exists } m \in E \text{ such that } f_i = f_{i\wedge m} > 0 \text{ for } i \in E\},$$
$$\widetilde{\mathscr{F}}_{I} = \big\{f : \text{ there exists } m \in E \text{ such that } f_i = f_{i\wedge m} > 0 \text{ for } i \in E \text{ and } f \text{ is strictly increasing in } \{1, \ldots, m\}\big\}.$$

Here, we use again the convention: $1/0 = \infty$. Note that for the localization, f is stopped at m rather than vanishing after m used in Sections 2 and 3. This is due

to the fact that the Neumann boundary is imposed at m but not the Dirichlet one. Besides, for the operator $I\!I$ here, the restriction on $\operatorname{supp}(f)$ used in Section 2 is no longer needed. Finally, define a local operator $\widetilde{R}$ (depending on m) acting on

$$\widetilde{\mathscr{V}}_1 = \cup_{m\in E}\big\{v: 1<v_i<1+a_i\big(1-v_{i-1}^{-1}\big)b_i^{-1} \text{ for } i=1,2,\ldots,m-1 \\ \text{and } v_i\!=\!1 \text{ for } i\geqslant m\big\}$$

by replacing a_m with $\tilde{a}_m := \mu_m a_m\big/\sum_{k=m}^N \mu_k$ in $R_i(v)$ for the same m as in $\widetilde{\mathscr{V}}_1$. Again, the change of a_m is due to the Neumann boundary at m. Note that if $v_i=1$ for all $i\geqslant m$, then $\widetilde{R}_i(v)=R_i(v)=0$ for all $i>m$.

Before stating our main results in this section, we mention an exceptional case that $\sum_i \mu_i=\infty$. On the one hand, by choosing $f_0=0$ and $f_i=1$ for $i\geqslant 1$, it follows that

$$D(f)=\mu_1 a_1<\infty, \qquad \mu(f^2)=\sum_{i\geqslant 1}\mu_i=\infty$$

and so $\lambda_0=0$. On the other hand, if $\sum_{i=1}^N \mu_i<\infty$, then for every f with $\mu(f^2)=\infty$, by setting $f^{(m)}=f_{\cdot\wedge m}\in L^2(\mu)$, we get

$$\infty > D\big(f^{(m)}\big)=\sum_{i=1}^m \mu_i a_i (f_i-f_{i-1})^2 \ \uparrow\ D(f) \qquad \text{as } m\to\infty,$$
$$\infty>\mu\big(f^{(m)\,2}\big)\geqslant \sum_{i=1}^m \mu_i f_i^2 \to\infty=\mu(f^2) \qquad \text{as } m\to\infty.$$

In words, for each non-square-integrable function f, both $\mu(f^2)$ and $D(f)$ can be approximated by a sequence of square-integrable ones. Hence, we can rewrite λ_0 as follows:

$$\lambda_0=\inf\{D(f):\mu(f^2)=1\}. \tag{4.2}$$

In this case, as will be seen soon but not obvious, we also have

$$\lambda_0=\inf\big\{D(f)\colon \mu(f^2)\!=\!1, f_i\!=\!f_{i\wedge m} \text{ for some } m\!\in\! E \text{ and all } i\!\in\! E\big\}, \tag{4.3}$$

Besides, we mention that the Dirichlet eigenvalue λ_0 is independent of $b_0\geqslant 0$ (cf. [4; Theorem 3.4] or [12; Theorem 3.7]).

For a large part of the paper, we do not use the uniqueness condition (1.2) (note that a change of a finite number of the rates a_i and b_i does not interfere in the uniqueness). Under (1.2), the process is ergodic iff $\sum_i \mu_i<\infty$ (see [10; Theorem 4.45 (2)], for instance). If (1.2) fails but $N=\infty$, then the decay rate for the minimal process is delayed to Section 7. In (2.2), the condition "$f\in\mathscr{K}$" means that we deal with the minimal process. This condition is removed in (4.2). It means that we are in this section dealing with the maximal process in the sense that the domain $\mathscr{D}^{\max}(D)$ of D ignored in (4.2) is taken to be the largest one: $\{f\in L^2(\mu): D(f)<\infty\}$ (that is the maximal process described at the beginning of Section 6 but killed at 1). When $N=\infty$, even though there is now a killing

at 1 (i.e., $a_1 > 0$), the regularity for (or the uniqueness of) the Dirichlet form is still equivalent to (1.3):

$$\sum_{k=1}^{\infty}\bigg(\frac{1}{b_k\mu_k}+\mu_k\bigg)=\infty \tag{1.3'}$$

since a modification of a finite number of rates does not change the regularity (cf. Theorem 9.22 for further information). In this section and Section 6, starting from any point in E, even though the process can visit every larger state, it will come back in a finite time. In this sense, the point infinity is regarded as a reflecting boundary.

It is the position to finish the comparison of (4.1) and (4.2). We have seen that $\lambda_0^{(4.1)} = \lambda_0^{(4.2)}$ once $\sum_i \mu_i < \infty$. We now claim that they can be different otherwise. To see this, note that on the one hand, $\lambda_0^{(4.1)} = 0$ if $\sum_i \mu_i = \infty$, as proved above. On the other hand, once (1.3)' holds $\big($in particular, if $\sum_i \mu_i = \infty$, then$\big)$ by Proposition 1.3, $\lambda_0^{(4.2)}$ coincides with

$$\inf\big\{D(f) : f \in \mathscr{K},\ \mu\big(f^2\big) = 1\big\},$$

which is the one used in (7.1) below and can often be non-zero. Thus, in general, $\lambda_0^{(4.2)} \geqslant \lambda_0^{(4.1)}$ and they can be different. As will be seen in Theorem 7.1 (2), in the special case that both of the series in (1.3)' are divergent, we have $\lambda_0^{(4.2)} = \lambda_0^{(7.1)} = 0$.

Theorem 4.1. *Assume that $\sum_{i=1}^N \mu_i < \infty$. Then the following variational formulas hold for λ_0 defined by one of (4.1)—(4.3).*

(1) *Difference form:*

$$\inf_{v\in\widetilde{\mathscr{V}}_1}\sup_{i\in E}\widetilde{R}_i(v)=\lambda_0=\sup_{v\in\mathscr{V}_1}\inf_{i\in E}R_i(v).$$

(2) *Single summation form:*

$$\inf_{f\in\widetilde{\mathscr{F}}_I}\sup_{i\in E}I_i(f)^{-1}=\lambda_0=\sup_{f\in\mathscr{F}_I}\inf_{i\in E}I_i(f)^{-1},$$

(3) *Double summation form:*

$$\lambda_0=\sup_{f\in\mathscr{F}_{II}}\inf_{i\in E}I\!I_i(f)^{-1}=\sup_{f\in\mathscr{F}_I}\inf_{i\in E}I\!I_i(f)^{-1},$$
$$\lambda_0=\inf_{f\in\widetilde{\mathscr{F}}_I}\sup_{i\in E}I\!I_i(f)^{-1}=\inf_{f\in\widetilde{\mathscr{F}}_{II}}\sup_{i\in E}I\!I_i(f)^{-1}=\inf_{f\in\widetilde{\mathscr{F}}_{II}\cup\widetilde{\mathscr{F}}'_{II}}\sup_{i\in E}I\!I_i(f)^{-1},$$

where $\widetilde{\mathscr{F}}'_{II}=\big\{f : f_i > 0 \text{ for all } i\in E \text{ and } fI\!I(f)\in L^2(\mu)\big\}.$

The next result was proved in [6] except the exceptional case that $\sum_i \mu_i = \infty$ in which case $\lambda_0 = 0$ (and $\delta = \infty$) and so the assertion is trivial. See also Corollary 5.2 below. Note that (ν_j) below is different from (2.15).

Theorem 4.2 (Criterion and basic estimates). *The rate* λ_0 *defined by (4.1)* $\big($*or equivalently by (4.2) provided* $\sum_{i\in E}\mu_i<\infty\big)$ *is positive iff* $\delta<\infty$*, where*

$$\delta=\sup_{n\in E}\nu[1,n]\,\mu[n,N]=\sup_{n\in E}\sum_{i=1}^{n}\frac{1}{\mu_i a_i}\sum_{j=n}^{N}\mu_j. \tag{4.4}$$

More precisely, we have $(4\delta)^{-1}\leqslant\lambda_0\leqslant\delta^{-1}$. *In particular, we have* $\lambda_0=0$ *if* $\sum_{i\in E}\mu_i=\infty$ *and* $\lambda_0>0$ *if either* $N<\infty$ *or (1.3)' fails.*

Theorem 4.3 (Approximating procedure). *Assume that* $\sum_{i=1}^{N}\mu_i<\infty$ *and* $\delta<\infty$. *Write* $\varphi_0=0$, $\varphi_i=\nu[1,i]:=\sum_{j=1}^{i}(\mu_j a_j)^{-1}$, $i\in E$.

(1) *Define* $f_1=\sqrt{\varphi}$, $f_n=f_{n-1}II(f_{n-1})$ *and* $\delta_n=\sup_{i\in E}II_i(f_n)$. *Then* δ_n *is decreasing in* n *and*

$$\lambda_0\geqslant\delta_\infty^{-1}\geqslant\cdots\geqslant{\delta_1}^{-1}\geqslant(4\delta)^{-1}.$$

(2) *For fixed* $m\in E$, *define*

$$\begin{aligned}f_1^{(m)}&=\varphi(\cdot\wedge m),\\ f_n^{(m)}&=\big[f_{n-1}^{(m)}II\big(f_{n-1}^{(m)}\big)\big](\cdot\wedge m),\qquad n\geqslant 2\end{aligned}$$

and then define $\delta_n'=\sup_{m\in E}\inf_{i\in E}II_i\big(f_n^{(m)}\big)$. *Then* δ_n' *is increasing in* n *and*

$$\delta^{-1}\geqslant{\delta_1'}^{-1}\geqslant\cdots\geqslant{\delta_\infty'}^{-1}\geqslant\lambda_0.$$

Next, define

$$\bar\delta_n=\sup_{m\in E}\frac{\mu\big(f_n^{(m)\,2}\big)}{D\big(f_n^{(m)}\big)},\qquad n\in E.$$

Then $\bar\delta_n^{-1}\geqslant\lambda_0$, $\bar\delta_{n+1}\geqslant\delta_n'$ *for all* $n\geqslant1$ *and* $\bar\delta_1=\delta_1'$.

As the first step given in Theorem 4.3, we obtain the following improvement of Theorem 4.2.

Corollary 4.4 (Improved estimates). *For the rate* λ_0 *defined by (4.1)* $\big($*or equivalently by (4.2) provided* $\sum_{i\in E}\mu_i<\infty\big)$, *we have*

$$\delta^{-1}\geqslant{\delta_1'}^{-1}\geqslant\lambda_0\geqslant\delta_1^{-1}\geqslant(4\delta)^{-1},$$

where

$$\begin{aligned}\delta_1&=\sup_{i\in E}\frac{1}{\sqrt{\varphi_i}}\sum_{k\geqslant1}\mu_k\varphi_{i\wedge k}\sqrt{\varphi_k}\\ &=\sup_{i\in E}\bigg[\frac{1}{\sqrt{\varphi_i}}\sum_{1\leqslant k<i}\mu_k\varphi_k^{3/2}+\sqrt{\varphi_i}\sum_{k=i}^{N}\mu_k\sqrt{\varphi_k}\bigg],\end{aligned} \tag{4.5}$$

$$\delta_1'=\sup_{m\in E}\frac{1}{\varphi_m}\sum_{k=1}^{N}\mu_k\varphi_{k\wedge m}^2=\sup_{m\in E}\bigg[\frac{1}{\varphi_m}\sum_{k=1}^{m-1}\mu_k\varphi_k^2+\varphi_m\mu[m,N]\bigg]\in[\delta,\,2\delta]. \tag{4.6}$$

Proof of Theorem 4.1. Note that

$$I_i(f) = \frac{1}{\mu_i a_i(f_i - f_{i-1})} \sum_{j=i}^{N} \mu_j f_j = \frac{1}{\mu_{i-1} b_{i-1}(f_i - f_{i-1})} \sum_{j=i}^{N} \mu_j f_j.$$

Hence, $I_i(f)$ coincides with $I_{i-1}(f)$ used in [3, 4, 6, 7, 12], whenever $b_0 > 0$. The same change is made for the operator $I\!I(f)$ in this section.

Throughout this proof, we use $\lambda_0 = \lambda_0^{(4.3)}$ to denote the one given in (4.3). Similar to the proofs of Theorem 2.4 and Proposition 2.5, we adopt the following circle arguments:

$$\begin{aligned}
\lambda_0 &\geqslant \lambda_0^{(4.2)} && (4.7)\\
&\geqslant \sup_{f\in \mathscr{F}_{I\!I}} \inf_{i\in E} I\!I_i(f)^{-1} = \sup_{f\in \mathscr{F}_{I}} \inf_{i\in E} I\!I_i(f)^{-1} = \sup_{f\in \mathscr{F}_{I}} \inf_{i\in E} I_i(f)^{-1} && (4.8)\\
&\geqslant \sup_{v\in \mathscr{V}_1} \inf_{i\in E} R_i(v) && (4.9)\\
&\geqslant \lambda_0 && (4.10)
\end{aligned}$$

and

$$\begin{aligned}
\lambda_0 &\leqslant \inf_{f\in \widetilde{\mathscr{F}}_{I\!I}\cup \widetilde{\mathscr{F}}'_{I\!I}} \sup_{i\in E} I\!I_i(f)^{-1} && (4.11)\\
&\leqslant \inf_{f\in \widetilde{\mathscr{F}}_{I\!I}} \sup_{i\in E} I\!I_i(f)^{-1} = \inf_{f\in \widetilde{\mathscr{F}}_{I}} \sup_{i\in E} I\!I_i(f)^{-1} = \inf_{f\in \widetilde{\mathscr{F}}_{I}} \sup_{i\in E} I_i(f)^{-1} && (4.12)\\
&\leqslant \inf_{v\in \widetilde{\mathscr{V}}_1} \sup_{i\in E} \widetilde{R}_i(v) && (4.13)\\
&\leqslant \lambda_0 && (4.14)
\end{aligned}$$

Assertion (4.7) is obvious. The following assertions are proved in [4; Theorem 3.3], or [12; §3.8] and [7; §2] (see also the remark given in the next paragraph):

$$\sup_{f\in \mathscr{F}_{I}} \inf_{i\in E} I_i(f)^{-1} = \sup_{f\in \mathscr{F}_{I}} \inf_{i\in E} I\!I_i(f)^{-1} = \sup_{f\in \mathscr{F}_{I\!I}} \inf_{i\in E} I\!I_i(f)^{-1} \leqslant \lambda_0^{(4.2)}. \tag{4.15}$$

$$\inf_{f\in \widetilde{\mathscr{F}}_{I}} \sup_{i\in E} I_i(f)^{-1} = \inf_{f\in \widetilde{\mathscr{F}}_{I}} \sup_{i\in E} I\!I_i(f)^{-1} = \inf_{f\in \widetilde{\mathscr{F}}_{I\!I}} \sup_{i\in E} I\!I_i(f)^{-1}. \tag{4.16}$$

In particular, we have known (4.8) and (4.12) since the inequality in (4.12) is trivial. It remains to prove (4.9)–(4.11), (4.13) and (4.14).

In [7; §2] and [12; §3.8], only the ergodic case under condition (1.2) is considered. But for (4.15) and (4.16), one does not need (1.2). Actually, one can now follow the proofs of Theorem 2.4 and Proposition 2.5 with a little change. For instance, to prove the last inequality in (4.15), following proof (a) of Theorem 2.4 and Proposition 2.5, let g satisfy $\|g\| = 1$ and $g_0 = 0$. Then

$$\begin{aligned}
1 &= \sum_{i\in E} \mu_i g_i^2 \quad (\text{since } \|g\| = 1)\\
&= \sum_{i\in E} \mu_i \bigg(\sum_{j=1}^{i} (g_j - g_{j-1})\bigg)^2 \quad (\text{since } g_0 = 0)\\
&\leqslant \sum_{i\in E} \mu_i \sum_{j=1}^{i} \frac{(g_j - g_{j-1})^2 \mu_j a_j}{h_j} \sum_{k=1}^{i} \frac{h_k}{\mu_k a_k}.
\end{aligned}$$

Exchanging the order of the first two sums on the right-hand side, we get

$$\begin{aligned}1 &\leqslant \sum_{j\in E}\mu_j a_j (g_j - g_{j-1})^2 \frac{1}{h_j}\sum_{i=j}^N \mu_i \sum_{k=1}^i \frac{h_k}{\mu_k a_k}\\ &\leqslant D(g)\sup_{j\in E}\frac{1}{h_j}\sum_{i=j}^N \mu_i \sum_{k=1}^i \frac{h_k}{\mu_k a_k}\\ &=: D(g)\sup_{j\in E} H_j.\end{aligned}$$

The next step is to choose $h_j = \sum_{i=j}^N \mu_i f_i$ for a given $f \in \mathscr{F}_{II}$ with $\sup_{j\in E} II_j(f) < \infty$. From these, it should be clear what change is required in order to prove (4.15) and (4.16).

We now begin to work on the additional part of the proof.

(a) Prove that $\sup_{f\in\mathscr{F}_{II}} \inf_{i\in E} II_i(f)^{-1} \geqslant \sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v)$.

As in proof (c) of Theorem 2.4 and Proposition 2.5, we use $R_i(u)$,

$$R_i(u) = a_i\bigg(1 - \frac{u_{i-1}}{u_i}\bigg) + b_i\bigg(1 - \frac{u_{i+1}}{u_i}\bigg), \qquad i \in E,$$

(u_{N+1} is free if $N < \infty$ since $b_N = 0$), instead of $R_i(v)$, where $u_i > 0$ for $i \in E$ and $u_0 = 0$. Then $v_i > 1\,(i \in E)$ means that $u_{i+1} > u_i > 0$, and $v_i = 1$ for $i \geqslant m$ means that $u_i = u_{i\wedge m} > 0$.

Without loss of generality, assume that $\inf_{i\in E} R_i(u) > 0$ for a given strictly increasing u with $u_0 = 0$. Define $f_i = (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1} \big[= u_i R_i(u)\big]$ for $i \in E$ and $f_0 = 0$. Then by assumption,

$$f_i/u_i = R_i(u) > 0, \qquad i \in E.$$

Hence, $f \in \mathscr{F}_{II}$. Next, since

$$0 < \mu_k f_k = \mu_k a_k (u_k - u_{k-1}) - \mu_{k+1} a_{k+1}(u_{k+1} - u_k)$$

and the strictly increasing property of u_i in i, it follows that

$$0 < \sum_{k=j}^N \mu_k f_k \leqslant \mu_j a_j (u_j - u_{j-1}),$$

and so

$$u_i = \sum_{j=1}^i (u_j - u_{j-1}) \geqslant \sum_{j=1}^i \nu_j \sum_{k=j}^N \mu_k f_k > 0.$$

We obtain

$$R_i(u) = f_i/u_i \leqslant II_i(f)^{-1}, \qquad i \in E.$$

Therefore, we have first

$$\inf_{i\in E} R_i(u) \leqslant \inf_{i\in E} I\!I_i(f)^{-1} \leqslant \sup_{f\in\mathscr{F}_{I\!I}} \inf_{i\in E} I\!I_i(f)^{-1},$$

and then

$$\sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \leqslant \sup_{f\in\mathscr{F}_{I\!I}} \inf_{i\in E} I\!I_i(f)^{-1},$$

as required.

(b) Prove that $\sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \geqslant \lambda_0$.

First, we show that $\sup_{v\in\mathscr{V}_1} \inf_{i\in E} R_i(v) \geqslant 0$. For a given positive $f \in L^1(\mu)$, let $u = fI\!I(f)$. Then $u_{i+1}/u_i > 1$ and $R_i(u) = f_i/u_i > 0$ for all $i \in E$. With $(v_i = u_{i+1}/u_i) \in \mathscr{V}_1$, this implies $\inf_{i\in E} R_i(v) \geqslant 0$ and then the required assertion follows.

Alternatively, since $a_1 > 0$, the eigenfunction is still strictly increasing when $\lambda_0 = 0$ by part (3) of Proposition 2.1. Hence the proof in the case of $\lambda_0 = 0$ can be combined into the next paragraph, and then the last paragraph can be omitted.

By assumption, we have $\sum_{i\in E} \mu_i < \infty$. When $\lambda_0 > 0$, it was proved in proof (d) of [12; Theorem 3.7] that the eigenfunction of $\lambda_0^{(4.2)}$ is strictly increasing. Even though λ_0 could formally be bigger than $\lambda_0^{(4.2)}$, the same proof still works for the eigenfunction g of λ_0 since the modified function $\bar g$ used there satisfies $\bar g_i = \bar g_{i\wedge n}$ for some n. Having this at hand, the proof is just a use of the eigenequation:

$$-\Omega g(i) := -b_i(g_{i+1} - g_i) + a_i(g_i - g_{i-1}) = \lambda_0 g_i, \qquad i \in E,\ g_0 := 0$$

(g_{N+1} is free if $N < \infty$ since $b_N = 0$). With $v_i := g_{i+1}/g_i > 1$ for $i < N$, this gives us $v \in \mathscr{V}_1$ and $R_i(v) \equiv \lambda_0$, and so the assertion follows.

We have thus completed the circle argument of (4.7)—(4.10).

(c) Prove that $\lambda_0 \leqslant \inf_{f\in\widetilde{\mathscr{F}}_{I\!I}\cup\widetilde{\mathscr{F}}'_{I\!I}} \sup_{i\in E} I\!I_i(f)^{-1}$.

In the original proof of [7; Theorem 2.1], when $N = \infty$, from the estimate

$$D(g) \leqslant \mu(g^2) \sup_{i\in E} I\!I_i(f)^{-1}$$

for $f \in \widetilde{\mathscr{F}}_{I\!I}$ and $g := [fI\!I(f)](\cdot \wedge m)$ to conclude that $\lambda_0 \leqslant D(g)/\mu(g^2)$, one requires an additional condition $g \in L^2(\mu)$, provided $m = \infty$ is allowed. This is the reason why the set $\widetilde{\mathscr{F}}'_{I\!I}$ in part (3) of Theorem 4.1 is added. Anyhow, with the modified conditions, the same proof gives us the required assertion (cf. proof (f) of Theorem 2.4 and Proposition 2.5).

(d) Prove that $\inf_{f\in\widetilde{\mathscr{F}}_{I\!I}} \sup_{i\in E} I\!I_i(f)^{-1} \leqslant \inf_{v\in\widetilde{\mathscr{V}}_1} \sup_{i\in E} \widetilde{R}_i(v)$.

Given u with $u_0 = 0$ and $u_i = u_{i\wedge m}$ for all $i \in E$ so that $(v_i := u_{i+1}/u_i) \in \widetilde{\mathscr{V}}_1$, let

$$f_i = \begin{cases} (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}, & i \leqslant m-1 \\ \tilde a_m (u_m - u_{m-1}), & i \geqslant m. \end{cases}$$

It is simple to check that $f_0 = 0$,

$$f_i/u_i = \widetilde{R}_i(u) > 0 \text{ for } i \in \{1, \dots, m\} \text{ and } f_i = f_m \text{ for } i > m,$$

and so $f \in \widetilde{\mathscr{F}}_{II}$. Moreover, since

$$\begin{aligned}\sum_{k=j}^{m-1} \mu_k f_k &= \mu_j a_j (u_j - u_{j-1}) - \mu_m a_m (u_m - u_{m-1}) \\ &= \mu_j a_j (u_j - u_{j-1}) - f_m \sum_{k=m}^{N} \mu_k,\end{aligned}$$

we get

$$0 < \sum_{k=j}^{N} \mu_k f_k = \mu_j a_j (u_j - u_{j-1}).$$

It follows that

$$0 < u_i = \sum_{j=1}^{i} (u_j - u_{j-1}) = \sum_{j=1}^{i} \nu_j \sum_{k=j}^{N} \mu_k f_k, \qquad i \in \{1, \dots, m\},$$

and then $\widetilde{R}_i(u) = f_i/u_i = II_i(f)^{-1}$ for $i \in \{1, 2, \dots, m\}$. Therefore, we have

$$\max_{1 \leqslant i \leqslant m} \widetilde{R}_i(u) = \max_{1 \leqslant i \leqslant m} II_i(f)^{-1} \geqslant \inf_{f \in \widetilde{\mathscr{F}}_{II},\, f_i = f_{i \wedge m}} \max_{1 \leqslant i \leqslant m} II_i(f)^{-1} \geqslant \inf_{f \in \widetilde{\mathscr{F}}_{II}} \sup_{i \in E} II_i(f)^{-1},$$

and then

$$\inf_{v \in \widetilde{\mathscr{V}}_1} \sup_{i \in E} \widetilde{R}_i(v) \geqslant \inf_{f \in \widetilde{\mathscr{F}}_{II}} \sup_{i \in E} II_i(f)^{-1}.$$

(e) Prove that $\inf_{v \in \widetilde{\mathscr{V}}_1} \sup_{i \in E} \widetilde{R}_i(v) \leqslant \lambda_0$.

Recall the definition of λ_0:

$$\lambda_0 = \inf \big\{ D(f) : \mu(f^2) = 1, f_i = f_{i \wedge m} \text{ for some } m \in E \text{ and all } i \in E \big\}.$$

Clearly, we have

$$\lambda_0^{(m)} := \inf \big\{ D(f) : \mu(f^2) = 1,\ f_i = f_{i \wedge m} \text{ for all } i \in E \big\} \downarrow \lambda_0 \ \text{ as } m \uparrow N.$$

We now explain the meaning of $\lambda_0^{(m)}$ as follows. Let

$$\begin{gathered}\tilde{\mu}_i = \mu_i,\ 1 \leqslant i < m, \quad \tilde{\mu}_m = \sum_{i=m}^{N} \mu_i, \\ \tilde{a}_i = a_i,\ 1 \leqslant i < m, \quad \tilde{a}_m = \mu_m a_m / \tilde{\mu}_m, \\ \widetilde{D}(f) = \sum_{i=1}^{m} \tilde{\mu}_i \tilde{a}_i (f_i - f_{i-1})^2. \end{gathered} \tag{4.17}$$

Then $\tilde{\mu}_i\tilde{a}_i = \mu_i a_i$ for $i = 1, \dots, m$, $\widetilde{D}(f) = D(f)$ and $\tilde{\mu}(f^2) = \mu(f^2)$ for every f with $f = f_{\cdot\wedge m}$. Thus, $\lambda_0^{(m)}$ is just the first eigenvalue of the local Dirichlet form $\big(\widetilde{D}, \mathscr{D}\big(\widetilde{D}\big)\big)$ having the state space $\{1, \dots, m\}$, with Dirichlet (absorbing) boundary at 0 and Neumann (reflecting) boundary at m. Let g ($g_0 = 0$) be the eigenfunction of the local first eigenvalue $\lambda_0^{(m)}$. Extend g to the whole space by setting $g_i = g_{i\wedge m}$. Next, set $u_i = g_i$ for $i < N$. Then

$$\widetilde{R}_i(u) = \begin{cases} \lambda_0^{(m)} > 0, & i \in \{1, \dots, m\}, \\ 0, & i > m. \end{cases} \tag{4.18}$$

Furthermore, for $v_i := u_{i+1}/u_i$, we have $v_0 = \infty$, $v_i > 1$ on $\{1, \dots, m-1\}$, and $v_i = 1$ for $i \geqslant m$. Thus, by (4.18), it is easy to check that $v = (v_i) \in \widetilde{\mathscr{V}}_1$. Therefore,

$$\begin{aligned} \lambda_0^{(m)} &= \max_{1\leqslant i\leqslant m} \widetilde{R}_i(v) \\ &\geqslant \inf_{v\in\widetilde{\mathscr{V}}_1:\, v_i=1 \text{ for } i\geqslant m} \max_{1\leqslant i\leqslant m} \widetilde{R}_i(v) \\ &\geqslant \inf_{v\in\widetilde{\mathscr{V}}_1:\, v_i=1 \text{ for } i\geqslant \text{ some } n>1} \sup_{i\in E} \widetilde{R}_i(v) \\ &= \inf_{v\in\widetilde{\mathscr{V}}_1} \sup_{i\in E} \widetilde{R}_i(v). \end{aligned}$$

The assertion now follows by letting $m \to N$. $\square$

Proof of Theorem 4.3.

(a) We remark that the sequence $\big\{f_n^{(m)}\big\}_{n\in E}$ is clearly contained in $\widetilde{\mathscr{F}_I}$. But the modified sequence used in [7; Theorem 2.2],

$$\tilde{f}_1^{(m)} = \varphi(\cdot\wedge m), \qquad \tilde{f}_n^{(m)} = \tilde{f}_{n-1}^{(m)}(\cdot\wedge m)\, I\!I\big(\tilde{f}_{n-1}^{(m)}(\cdot\wedge m)\big), \quad n \geqslant 2,$$

is usually not contained in $\widetilde{\mathscr{F}_{I\!I}}$. However,

$$\begin{aligned} \delta_n' &= \sup_{m\in E}\inf_{i\in E} I\!I_i\big(f_n^{(m)}\big) \\ &= \sup_{m\in E}\min_{1\leqslant i\leqslant m} I\!I_i\big(f_n^{(m)}\big) \\ &= \sup_{m\in E}\min_{1\leqslant i\leqslant m} I\!I_i\big(\tilde{f}_n^{(m)}(\cdot\wedge m)\big) \\ &= \sup_{m\in E}\inf_{i\in E} I\!I_i\big(\tilde{f}_n^{(m)}(\cdot\wedge m)\big). \end{aligned}$$

Here in the last step, we have used the convention $1/0 = \infty$. Hence, these two sequences produce the same $\{\delta_n'\}$.

(b) The approximating procedure given in Theorem 4.3 is mainly a copy of [7; Theorem 2.2] (cf. the proof of Theorem 3.2). For later use, here we review the proof of part (1). From [6; proof of Theorem 3.5], we have known that

$$I_j(f_1) = \frac{1}{\mu_j a_j (f_1(j) - f_1(j-1))} \sum_{k\geqslant j} \mu_k f_1(k) \leqslant 4\delta, \qquad j \geqslant 1.$$

Hence (Alternatively, by the proportional property),

$$f_2(i) = \sum_{j=1}^{i} \frac{1}{\mu_j a_j} \sum_{k \geqslant j} \mu_k f_1(k) \leqslant 4\delta \sum_{j=1}^{i} (f_1(j) - f_1(j-1)) = 4\delta f_1(i).$$

This gives us the assertion $\delta_1 = \sup_{i \geqslant 1} I\!I_i(f_1) \leqslant 4\delta$.

To prove the monotonicity of $\{\delta_n\}$ and $\{f_n\} \subset L^1(\mu)$, we adopt induction. As we have just seen,

$$\delta_1 = \sup_{i \geqslant 1} \frac{f_2(i)}{f_1(i)} \leqslant 4\delta.$$

This means that $f_1 \in L^1(\mu)$ (or equivalently, $f_2 < \infty$) and $\delta_1 < \infty$ since $\delta < \infty$ by assumption. Assume that $f_n \in L^1(\mu)$ (or equivalently, $f_{n+1} < \infty$) and $\delta_n < \infty$. Then

$$\sum_{k \geqslant j} \mu_k f_{n+1}(k) = \sum_{k \geqslant j} \mu_k f_n(k)[f_{n+1}(k)/f_n(k)] \leqslant \delta_n \sum_{k \geqslant j} \mu_k f_n(k).$$

Multiplying both sides by ν_j and making summation from 1 to i, it follows that

$$f_{n+2}(i) \leqslant \delta_n f_{n+1}(i), \qquad i \geqslant 1.$$

Since $f_{n+1} < \infty$ and $\delta_n < \infty$ by assumption, we have $f_{n+2} < \infty$, and

$$\frac{f_{n+2}(i)}{f_{n+1}(i)} = I\!I_i(f_{n+1}) < \infty, \qquad i \geqslant 1.$$

This proves not only $f_{n+1} \in L^1(\mu)$ but also $\delta_{n+1} \leqslant \delta_n < \infty$.

The assertion that $\bar{\delta}_n^{-1} \geqslant \lambda_0$ is obvious by (4.2). Similar to proof (b) of Theorem 3.2, the assertion $\bar{\delta}_{n+1} \geqslant \delta_n'$ is a consequence of the last part of the proof of [7; Theorem 2.1]. $\square$

Proof of Corollary 4.4.

(a) The degenerated case that $\sum_i \mu_i = \infty$ is trivial since $\lambda_0^{(4.1)} = 0$ and $\delta = \delta_1 = \delta_1' = \infty$. The main assertion of Corollary 4.4 is a consequence of Theorem 4.3. Here, we consider (4.6) only since the proof of (4.5) is easier. Note that

$$I\!I_i\big(f_1^{(m)}\big) = \frac{1}{\varphi_{i \wedge m}} \sum_{j=1}^{i} \frac{1}{\mu_j a_j} \sum_{k=j}^{N} \mu_k \varphi_{k \wedge m}.$$

The right-hand side is clearly increasing in i for $i \geqslant m$ and is decreasing (not hard to check) in i when $i \leqslant m$. Hence, $I\!I_i\big(f_1^{(m)}\big)$ achieves its minimum at $i = m$. Then, by exchanging the order of the sums, it follows that the minimum is equal to

$$\frac{1}{\varphi_m} \sum_{k=1}^{N} \mu_k \varphi_{k \wedge m}^2.$$

This observation is due to Sirl, Zhang and Pollett (2007). We have thus proved the first equality in (4.6).

Next, following the proof of [6; Theorem 3.5], we have

$$D\big(f_1^{(m)}\big) = \sum_{i=1}^{m} \mu_i a_i (\varphi_i - \varphi_{i-1})^2 = \varphi_m,$$

and

$$\mu\big(f_1^{(m)\,2}\big) = \sum_{k=1}^{N} \mu_k \varphi_{k\wedge m}^2.$$

Combining these facts together, it follows that $\bar{\delta}_1 = \delta_1'$.

(b) Finally, we prove the estimates in (4.6). The lower estimate of δ_1' is rather easy since

$$\frac{1}{\varphi_m} \sum_{k=1}^{N} \mu_k \varphi_{k\wedge m}^2 \geqslant \frac{1}{\varphi_m} \sum_{k=m}^{N} \mu_k \varphi_{k\wedge m}^2 = \varphi_m \sum_{k=m}^{N} \mu_k.$$

For the upper estimate, use the summation by parts formula:

$$\sum_{k=1}^{N} \mu_k \varphi_{k\wedge m}^2 = \sum_{k=1}^{m} \big[\varphi_k^2 - \varphi_{k-1}^2\big] \sum_{j=k}^{N} \mu_j = \sum_{k=1}^{m} \frac{\varphi_k + \varphi_{k-1}}{\mu_k a_k} \sum_{j=k}^{N} \mu_j.$$

It follows that

$$\frac{1}{\varphi_m} \sum_{k=1}^{N} \mu_k \varphi_{k\wedge m}^2 < \frac{2}{\varphi_m} \sum_{k=1}^{m} \frac{1}{\mu_k a_k} \bigg[\varphi_k \sum_{j=k}^{N} \mu_j\bigg] \leqslant \frac{2\delta}{\varphi_m} \sum_{k=1}^{m} \frac{1}{\mu_k a_k} = 2\delta.$$

The estimate now follows by making the supremum with respect to $m \in E$. $\square$

5. Dual approach

This section is devoted to the duality of the processes studied in the previous sections, as well as a duality to be used in the next two sections. Again, the section is ended by a class of examples.

Suppose that we are given a birth–death process with state space $E = \{i : 0 \leqslant i < N+1\}$ $(N \leqslant \infty)$, birth rates $b_i > 0$ ($b_0 > 0$, especially) but $b_N \geqslant 0$ if $N < \infty$, and death rates $a_i > 0$ but $a_0 = 0$. The case that $b_N > 0$ is used in this section while the case of $b_N = 0$ is for use in Section 7. Define a dual chain with state space $\widehat{E} = \{i : 1 \leqslant i < N' + 1\}$ and with rates as follows:

$$\hat{b}_0 = 0, \quad \hat{b}_i = a_i, \quad \hat{a}_i = b_{i-1}, \qquad i \in \widehat{E}, \tag{5.1}$$

where $a_{N+1} = b_{N+1} = 0$ if $N < \infty$ by convention and

$$N' = \begin{cases} N, & N < \infty \text{ and } b_N = 0, \\ N+1, & N < \infty \text{ and } b_N > 0, \\ \infty, & N = \infty. \end{cases}$$

The dual process with rates $(\hat{a}_i, \hat{b}_i)$ has an absorbing at 0. When $N < \infty$, for the dual process, the state $N+1$ is absorbing if $b_N = 0$ (then $\hat{a}_{N+1} = 0$ but $\hat{b}_N > 0$); otherwise, it is a reflecting boundary since $\hat{a}_{N+1} = b_N > 0$. In a word, the absorbing boundary is dual to the reflecting one and vice versa. This dual technique goes back to Karlin and McGregor (1957b, §6). Next, define

$$\hat{\mu}_1 = 1, \qquad \hat{\mu}_n = \frac{\hat{b}_1 \cdots \hat{b}_{n-1}}{\hat{a}_2 \cdots \hat{a}_n}, \qquad 2 \leqslant n < N' + 1. \tag{5.2}$$

When $N < \infty$ and $b_N > 0$, then $\hat{a}_{N+1} > 0$, and so $\hat{\mu}_n$ can be defined up to $n = N+1$. Otherwise, it can be defined up to $n = N$ only. It is now easy to check (noticing the difference of (ν_j) and $(\hat{\nu}_j)$) that

$$\hat{\mu}_n = \frac{b_0}{\mu_n a_n} = b_0 \nu_{n-1}, \quad \hat{\nu}_n := \frac{1}{\hat{\mu}_n \hat{a}_n} = \frac{1}{b_0}\mu_{n-1}, \qquad 1 \leqslant n < N' + 1. \tag{5.3}$$

Actually, the rates $(\hat{a}_i, \hat{b}_i)$ in (5.1) are determined by the transform given in (5.3): $\hat{\mu}_n = b_0 \nu_{n-1}$ and $\hat{\nu}_n = \mu_{n-1}/b_0$. From this, it follows that

$$\begin{gathered}\mu_n = b_0 \hat{\nu}_{n+1} = \hat{a}_1 \hat{\nu}_{n+1}, \quad \nu_n = \frac{1}{b_0}\hat{\mu}_{n+1} = \frac{1}{\hat{a}_1}\hat{\mu}_{n+1}, \qquad 0 \leqslant n < N', \\ \mu_N = \hat{a}_1 \big(\hat{\mu}_N \hat{b}_N\big)^{-1} \qquad \text{if } N < \infty \text{ and } b_N = 0,\end{gathered} \tag{5.4}$$

and so

$$\begin{aligned}\sum_{n=1}^{N'} \frac{1}{\hat{\mu}_n \hat{a}_n} &= \sum_{n=1}^{N'} \hat{\nu}_n = \frac{1}{b_0} \sum_{n=0}^{N'-1} \mu_n, \\ \sum_{n=1}^{N'} \hat{\mu}_n &= b_0 \sum_{n=1}^{N'} \nu_{n-1} = b_0 \sum_{n=0}^{N'-1} \frac{1}{\mu_n b_n}.\end{aligned} \tag{5.5}$$

Note that by (5.1),

$$a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1} v_i = \hat{b}_{i+1} + \hat{a}_{i+1} - \frac{\hat{b}_i}{v_{i-1}} - \hat{a}_{i+2}\, v_i.$$

By a change of the variables $(v_i) \in \mathscr{V}$:

$$v_i = \frac{\hat{b}_{i+1}}{\hat{a}_{i+2}}\, \hat{v}_{i+1}, \tag{5.6}$$

or

$$\hat{v}_i = \frac{\hat{a}_{i+1}}{\hat{b}_i}\, v_{i-1} = \frac{b_i}{a_i}\, v_{i-1}, \tag{5.7}$$

we get

$$a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1} v_i = \hat{a}_{i+1}\left(1 - \frac{1}{\hat{v}_i}\right) + \hat{b}_{i+1}(1 - \hat{v}_{i+1}).$$

Since $b_0 > 0$, $v_{-1} > 0$ but $a_0 = 0$, from (5.7), it is clear that we should set $\hat{v}_0 = \infty$. Next, by (5.7) again,

$$v_{i-1} > \frac{a_i}{b_i} \Longleftrightarrow \hat{v}_i > 1.$$

It remains to examine the boundary condition on the right-hand side when $N < \infty$.

(1) First, let $b_N = 0$. Then $v = (v_i > 0 : 0 \leqslant i < N - 1)$, v_{-1} and v_{N-1} are free. The dual state space is $\widehat{E} = \{1, 2, \ldots, N\}$. The dual test sequence is $\hat{v} = \big(\hat{v}_i > 0 : 1 \leqslant i < N\big)$, $\hat{v}_N = 0$.

(2) Next, let $b_N > 0$. Then $v = (v_i > 0 : 0 \leqslant i < N)$, v_{-1} and v_N are free. The dual state space is $\widehat{E} = \{1, 2, \ldots, N+1\}$ with reflecting at $N + 1$. Hence, $\hat{v} = \big(\hat{v}_i > 0 : 1 \leqslant i < N + 1\big)$, $\hat{v}_{N+1} = 0$.

We have thus proved the following result.

Proposition 5.1. *For the dual processes defined above, the following identities hold:*

$$\begin{aligned}
&\sup_v \inf_{0 \leqslant i < N'} \left[a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1} v_i\right] \\
&\qquad = \sup_{\hat{v}} \inf_{1 \leqslant i < N'+1} \left[\hat{a}_i \left(1 - \frac{1}{\hat{v}_{i-1}}\right) + \hat{b}_i (1 - \hat{v}_i)\right],
\end{aligned} \tag{5.8}$$

where $v = (v_i > 0 : 0 \leqslant i < N' - 1)$ with free v_{-1}, and $\hat{v} = \big(\hat{v}_i > 0 : 1 \leqslant i < N'\big)$ with $\hat{v}_0 = \infty$, $v_{N'-1}$ is free and $\hat{v}_{N'} = 0$ if $N < \infty$;

$$\begin{aligned}
&\sup_{v \in \mathscr{V}_*} \inf_{0 \leqslant i < N+1} \left[a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1} v_i\right] \\
&\qquad = \sup_{\hat{v} \in \mathscr{V}_1} \inf_{1 \leqslant i < N+2} \left[\hat{a}_i \left(1 - \frac{1}{\hat{v}_{i-1}}\right) + \hat{b}_i (1 - \hat{v}_i)\right]
\end{aligned} \tag{5.9}$$

in the case that $b_N > 0$ if $N < \infty$, where $\mathscr{V}_*$ is given in Proposition 2.7, and $\mathscr{V}_1$ is defined in Theorem 4.1 replacing N by $N + 1$ when $N < \infty$.

In these formulas, $a_{N+1} = b_{N+1} = 0$ if $N < \infty$ by convention.

Corollary 5.2. *Given rates (a_i, b_i) as in Section 2 (then $b_N > 0$ if $N < \infty$), let $\lambda_0 = \lambda_0^{(2.2)}$ and define δ by (3.1). Next, define the dual rates $\big(\hat{a}_i, \hat{b}_i\big)$ as above. Correspondingly, we have $\hat{\lambda}_0$ and $\hat{\delta}$ defined by (4.1) and (4.4) replacing N by $N + 1$ if $N < \infty$, respectively, in terms of the dual rates. Then we have $\lambda_0 = \hat{\lambda}_0$ and $\delta = \hat{\delta}$.*

Proof. Having relationship (5.9) at hand, the assertion that $\lambda_0 = \hat{\lambda}_0$ follows by a combination part (2) of Proposition 2.7 and part (1) of Theorem 4.1, provided $\sum_i \hat{\mu}_i < \infty$.

Next, by (4.4), (5.3), and (3.1), we have

$$\begin{aligned}\hat{\delta} &= \sup_{1\leqslant i<N+2} \sum_{j=1}^{i} \frac{1}{\hat{\mu}_j \hat{a}_j} \sum_{j=i}^{N+1} \hat{\mu}_j \\ &= \sup_{1\leqslant i<N+2} \sum_{j=1}^{i} \hat{\nu}_j \sum_{k=i}^{N+1} \hat{\mu}_k \\ &= \sup_{1\leqslant i<N+2} \sum_{j=1}^{i} \frac{\mu_{j-1}}{b_0} \sum_{k=i}^{N+1} b_0 \nu_{k-1} \\ &= \sup_{0\leqslant i<N+1} \sum_{j=0}^{i} \mu_j \sum_{k=i}^{N} \frac{1}{\mu_k b_k} \\ &= \delta.\end{aligned}$$

This proves that $\delta = \hat{\delta}$. In particular, if $\sum_i \hat{\mu}_i\big(= \sum_i \nu_i\big) = \infty$, then by Theorem 3.1 and Corollary 4.4, we get $\lambda_0 = \hat{\lambda}_0 = 0$. We have thus completed the proof of $\lambda_0 = \hat{\lambda}_0$. $\square$

As will be seen in Theorem 7.1 (2), in the degenerated case that $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i b_i)^{-1} = \infty$, the dual of the process studied in Section 2 also goes to the one studied in Section 7.

Before moving further, let us discuss the duality used here. Very recently, Chi Zhang provides us a nice explanation which leads to a deeper understanding of the duality (5.1). Consider a simple example as follows:

$$Q = \begin{pmatrix} -b_0 & b_0 & 0 & 0 \\ a_1 & -a_1 - b_1 & b_1 & 0 \\ 0 & a_2 & -a_2 - b_2 & b_2 \\ 0 & 0 & a_3 & -a_3 - b_3 \end{pmatrix}, \qquad a_i, b_i > 0.$$

Introduce an invertible matrix:

$$M = \begin{pmatrix} \mu_0 b_0 & -\mu_0 b_0 & 0 & 0 \\ 0 & \mu_1 b_1 & -\mu_1 b_1 & 0 \\ 0 & 0 & \mu_2 b_2 & -\mu_2 b_2 \\ 0 & 0 & 0 & \mu_3 b_3 \end{pmatrix} \Longrightarrow M^{-1} = \begin{pmatrix} \frac{1}{\mu_0 b_0} & \frac{1}{\mu_1 b_1} & \frac{1}{\mu_2 b_2} & \frac{1}{\mu_3 b_3} \\ 0 & \frac{1}{\mu_1 b_1} & \frac{1}{\mu_2 b_2} & \frac{1}{\mu_3 b_3} \\ 0 & 0 & \frac{1}{\mu_2 b_2} & \frac{1}{\mu_3 b_3} \\ 0 & 0 & 0 & \frac{1}{\mu_3 b_3} \end{pmatrix}.$$

Then

$$\begin{aligned} MQM^{-1} &= \begin{pmatrix} -a_1 - b_0 & a_1 & 0 & 0 \\ b_1 & -a_2 - b_1 & a_2 & 0 \\ 0 & b_2 & -a_3 - b_2 & a_3 \\ 0 & 0 & b_3 & -b_3 \end{pmatrix} \\ &= \begin{pmatrix} -\hat{a}_1 - \hat{b}_1 & \hat{b}_1 & 0 & 0 \\ \hat{a}_2 & -\hat{a}_2 - \hat{b}_2 & \hat{b}_2 & 0 \\ 0 & \hat{a}_3 & -\hat{a}_3 - \hat{b}_3 & \hat{b}_3 \\ 0 & 0 & \hat{a}_4 & -\hat{a}_4 \end{pmatrix} \\ &= \widehat{Q}. \end{aligned}$$

Hence, the dual matrix $\widehat{Q}$ is just the classical similar transformation of Q and so they have the same spectrum. In particular, the eigenequation $Qg = -\lambda_0 g$ ($g \neq 0$) is transferred into

$$\widehat{Q}(Mg) = (MQM^{-1})(Mg) = \lambda_0 Mg = \hat{\lambda}_0\big(Mg\big).$$

Hence, the eigenfunction g of λ_0 is transformed to $\hat{g} = Mg$ of $\hat{\lambda}_0 = \lambda_0$. Correspondingly, the test function f is transformed to $\hat{f} = Mf$. From this, it should be clear that all the operators R and $\widehat{R}$, I and $\hat{I}$, $I\!I$ and $\widehat{I\!I}$ are closely related to each other and then so are the variational formulas.

Having these facts at hand, one can simplify a part of the previous proofs. However, we prefer to keep all the details here since they are needed when we go to the more general situation, so called the Poincaré-type inequalities (Section 8), or can be used as a reference for studying the continuous case. For the Poincaré-type inequalities, the current duality seems not available.

By Corollary 5.2, we have two ways to estimate $\lambda_0 = \hat{\lambda}_0$: using either the rates (a_i, b_i) or $\big(\hat{a}_i, \hat{b}_i\big)$. The corresponding formulas for δ_1', $\hat{\delta}_1'$, δ_1 and $\hat{\delta}_1$ are collected in Tables 5.1 and 5.2.

Table 5.1: Expressions of $\delta = \hat{\delta}$, δ_1', $\hat{\delta}_1'$, δ_1 and $\hat{\delta}_1$ in terms of the rates (b_i, a_i):

$$\delta = \hat{\delta} = \sup_{0\leqslant i<N+1} \mu[0,i]\,\nu[i,N] = \sup_{0\leqslant i<N+1} \sum_{j=0}^{i}\mu_j \sum_{k=i}^{N}\nu_k, \tag{5.10}$$

$$\begin{aligned}\delta_1' &= \sup_{0\leqslant i<N+1} \frac{1}{\nu[i,N]}\sum_{k=0}^{N}\mu_k\nu[k\vee i,N]^2\\ &= \sup_{0\leqslant i<N+1}\bigg[\mu[0,i]\nu[i,N] + \frac{1}{\nu[i,N]}\sum_{k=i+1}^{N}\mu_k\nu[k,N]^2\bigg],\end{aligned} \tag{5.11}$$

$$\begin{aligned}\hat{\delta}_1' &= \sup_{0\leqslant i<N+1} \frac{1}{\mu[0,i]}\sum_{k=0}^{N}\nu_k\mu[0,k\wedge i]^2\\ &= \sup_{0\leqslant i<N+1}\bigg[\mu[0,i]\nu[i,N] + \frac{1}{\mu[0,i]}\sum_{k=0}^{i-1}\nu_k\mu[0,k]^2\bigg],\end{aligned} \tag{5.12}$$

$$\begin{aligned}\delta_1 &= \sup_{0\leqslant i<N+1} \frac{1}{\sqrt{\nu[i,N]}}\sum_{k=0}^{N}\mu_k\nu[i\vee k,N]\sqrt{\nu[k,N]}\\ &= \sup_{0\leqslant i<N+1}\bigg[\sqrt{\nu[i,N]}\sum_{k=0}^{i}\mu_k\sqrt{\nu[k,N]} + \frac{1}{\sqrt{\nu[i,N]}}\sum_{k=i+1}^{N}\mu_k\nu[k,N]^{3/2}\bigg],\end{aligned} \tag{5.13}$$

$$\begin{aligned}\hat{\delta}_1 &= \sup_{0\leqslant i<N+1} \frac{1}{\sqrt{\mu[0,i]}}\sum_{k=0}^{N}\nu_k\mu[0,k\wedge i]\sqrt{\mu[0,k]}\\ &= \sup_{0\leqslant i<N+1}\bigg[\frac{1}{\sqrt{\mu[0,i]}}\sum_{k=0}^{i-1}\nu_k\mu[0,k]^{3/2} + \sqrt{\mu[0,i]}\sum_{k=i}^{N}\nu_k\sqrt{\mu[0,k]}\bigg].\end{aligned} \tag{5.14}$$

Table 5.2: Expressions of $\delta=\hat\delta$, δ_1', $\hat\delta_1'$, δ_1 and $\hat\delta_1$ in terms of the rates $(\hat b_i, \hat a_i)$:

$$\delta=\hat\delta=\sup_{1\leqslant i<N+1}\hat\nu[1,i]\,\hat\mu[i,N]=\sup_{1\leqslant i<N+1}\sum_{k=1}^{i}\hat\nu_k\sum_{j=i}^{N}\hat\mu_j, \tag{5.15}$$

$$\begin{aligned}\delta_1' &= \sup_{1\leqslant i<N+1}\frac{1}{\hat\mu[i,N]}\sum_{k=1}^{N}\hat\nu_k\hat\mu[k\vee i,N]^2\\ &=\sup_{1\leqslant i<N+1}\bigg[\hat\mu[i,N]\hat\nu[1,i]+\frac{1}{\hat\mu[i,N]}\sum_{k=i+1}^{N}\hat\nu_k\hat\mu[k,N]^2\bigg],\end{aligned} \tag{5.16}$$

$$\begin{aligned}\hat\delta_1' &= \sup_{1\leqslant i<N+1}\frac{1}{\hat\nu[1,i]}\sum_{k=1}^{N}\hat\mu_k\hat\nu[1,k\wedge i]^2\\ &=\sup_{1\leqslant i<N+1}\bigg[\hat\mu[i,N]\hat\nu[1,i]+\frac{1}{\hat\nu[1,i]}\sum_{k=1}^{i-1}\hat\mu_k\hat\nu[1,k]^2\bigg].\end{aligned} \tag{5.17}$$

$$\begin{aligned}\delta_1 &= \sup_{1\leqslant i<N+1}\frac{1}{\sqrt{\hat\mu[i,N]}}\sum_{k=1}^{N}\hat\nu_k\hat\mu[k\vee i,N]\sqrt{\hat\mu[k,N]}\\ &=\sup_{1\leqslant i<N+1}\bigg[\sqrt{\hat\mu[i,N]}\sum_{k=1}^{i}\hat\nu_k\sqrt{\hat\mu[k,N]}+\frac{1}{\sqrt{\hat\mu[i,N]}}\sum_{k=i+1}^{N}\hat\nu_k\,\hat\mu[k,N]^{3/2}\bigg],\end{aligned} \tag{5.18}$$

$$\begin{aligned}\hat\delta_1 &= \sup_{1\leqslant i<N+1}\frac{1}{\sqrt{\hat\nu[1,i]}}\sum_{k=1}^{N}\hat\mu_k\hat\nu[1,k\wedge i]\sqrt{\hat\nu[1,k]}\\ &=\sup_{1\leqslant i<N+1}\bigg[\frac{1}{\sqrt{\hat\nu[1,i]}}\sum_{k=1}^{i-1}\hat\mu_k\hat\nu[1,k]^{3/2}+\sqrt{\hat\nu[1,i]}\sum_{k=i}^{N}\hat\mu_k\sqrt{\hat\nu[1,k]}\bigg],\end{aligned} \tag{5.19}$$

The next four examples are dual of Examples 3.4–3.7, respectively.

Example 5.3. *For Example 3.4, we have $\hat a_i\equiv b\,(i\geqslant 1)$, $\hat b_i\equiv a\,(a>0)$, $b\geqslant a$. Then $\hat\delta=\delta=b(a-b)^{-2}$, $\hat\delta_1'=\delta_1'=(a+b)(a-b)^{-2}$, and $\hat\delta_1=\delta_1=\lambda_0^{-1}=\big(\sqrt{a}-\sqrt{b}\big)^{-2}$. In particular, if we take $\hat a_i=4$ and $\hat b_i=1\,(i\geqslant 1)$, then $\hat\lambda_0=1$,*

$$\begin{aligned}&\hat\delta_1'=5/9=0.\dot 5,\quad \hat\delta_2'=0.6\dot 4,\quad \hat\delta_3'\approx 0.71,\quad \hat\delta_4'\approx 0.755,\quad \hat\delta_5'\approx 0.79;\\ &\bar{\hat\delta}_1=0.\dot 5,\quad \bar{\hat\delta}_2\approx 0.71,\quad \bar{\hat\delta}_3\approx 0.79,\quad \bar{\hat\delta}_4\approx 0.835\quad \bar{\hat\delta}_5\approx 0.8647.\end{aligned}$$

Thus, $\hat\delta_n'$ and $\bar{\hat\delta}_n$ are increasing and close to $\hat\lambda_0^{-1}$ as $n\uparrow$.

Proof. To compute $\hat\delta_1'$ and $\hat\delta_1$, we use Table 5.1. For simplicity, write $\gamma=b/a>1$. Then

$$\mu_k=\gamma^k,\qquad \mu[0,i]=\frac{\gamma^{i+1}-1}{\gamma-1},\qquad \nu_k=\frac{1}{b}\gamma^{-k}.$$

(a) Note that

$$\begin{aligned}
&\frac{1}{\mu[0,i]}\sum_{k=0}^{i-1}\nu_k\mu[0,k]^2+\mu[0,i]\sum_{k=i}^{\infty}\nu_k\\
&=\frac{1}{b}\bigg[\frac{\gamma-1}{\gamma^{i+1}-1}\sum_{k=0}^{i-1}\gamma^{-k}\bigg(\frac{\gamma^{k+1}-1}{\gamma-1}\bigg)^2+\frac{\gamma^{i+1}-1}{\gamma-1}\sum_{k\geqslant i}\gamma^{-k}\bigg]\\
&=\frac{1}{b(\gamma-1)}\bigg[\frac{1}{\gamma^{i+1}-1}\sum_{k=0}^{i-1}\gamma^{-k}\big(\gamma^{k+1}-1\big)^2+\big(\gamma^{i+1}-1\big)\sum_{k\geqslant i}\gamma^{-k}\bigg]\\
&=\frac{1}{b(\gamma-1)}\bigg[\frac{\gamma(1+\gamma)}{\gamma-1}-\frac{2(i+1)\gamma}{\gamma^{i+1}-1}\bigg].
\end{aligned}$$

Since the second term in the last $[\cdots]$ is negative and $\gamma>1$, the right-hand side attains its supremum at $i=\infty$. By (5.12), we have thus obtained

$$\hat\delta_1'=\frac{\gamma(1+\gamma)}{b(\gamma-1)^2}=\frac{a+b}{(a-b)^2}.$$

(b) Next, note that

$$\begin{aligned}
&\frac{1}{\sqrt{\mu[0,i]}}\sum_{k=0}^{i-1}\nu_k\mu[0,k]^{3/2}+\sqrt{\mu[0,i]}\sum_{k=i}^{\infty}\nu_k\sqrt{\mu[0,k]}\\
&=\frac{1}{b(\gamma-1)}\bigg[\frac{1}{\sqrt{\gamma^{i+1}-1}}\sum_{k=0}^{i-1}\gamma^{-k}\big(\gamma^{k+1}-1\big)^{3/2}+\sqrt{\gamma^{i+1}-1}\sum_{k\geqslant i}\gamma^{-k}\sqrt{r^{k+1}-1}\bigg]\\
&\leqslant\frac{1}{b(\gamma-1)}\bigg[\frac{1}{\sqrt{\gamma^{i+1}-1}}\sum_{k=0}^{i-1}\gamma^{(k+3)/2}+\sqrt{\gamma^{i+1}-1}\sum_{k\geqslant i}\gamma^{-k/2+1/2}\bigg]\\
&=\frac{1}{b(\gamma-1)}\bigg[\frac{1}{\sqrt{\gamma^{i+1}-1}}\frac{\gamma^{3/2}(\gamma^{i/2}-1)}{\sqrt{\gamma}-1}+\frac{\gamma^{-i/2+1/2}\sqrt{\gamma^{i+1}-1}}{1-1/\sqrt{\gamma}}\bigg]\\
&\leqslant\frac{1}{b(\gamma-1)}\bigg[\frac{\gamma}{\sqrt{\gamma}-1}+\frac{\gamma\sqrt{\gamma}}{\sqrt{\gamma}-1}\bigg]\\
&=\frac{\gamma}{b(\sqrt{\gamma}-1)^2}\\
&=\frac{1}{(\sqrt{a}-\sqrt{b})^2}.
\end{aligned}$$

By (5.14), this means that $\hat\delta_1\leqslant\hat\lambda_0^{-1}$ and so the equality sign must hold because $\hat\delta_1^{-1}$ is a lower estimate: $\hat\lambda_0\geqslant\hat\delta_1^{-1}$.

(c) We now compute the approximating sequences $\{\hat\delta_n'\}$ and $\{\bar{\hat\delta}_n\}$ for the upper estimate, using the dual rate $(\hat a_i,\hat b_i)$. In the particular case, we have

$$\hat\mu_i=4^{1-i},\quad \hat\nu_i=4^{i-2},\quad \hat\varphi_i=\hat\nu[1,i]=\frac{4^i-1}{12}.$$

The approximating sequences can be computed successively by using the following formulas:

$$\begin{aligned}
f_1^{(m)}(i) &= \frac{4^i - 1}{12}, \qquad i \in \{1, 2, \dots, m\},\\
f_n^{(m)}(i) &= \frac{1}{3}\bigg\{\sum_{k=1}^{i-1}(1 - 4^{-k}) f_{n-1}^{(m)}(k) + (4^i - 1)\sum_{k=i}^{m-1} 4^{-k} f_{n-1}^{(m)}(k)\\
&\qquad + \frac{1}{3}(4^i - 1) 4^{1-m} f_{n-1}^{(m)}(m)\bigg\}, \qquad i \in \{1, 2, \dots, m\},\ n \geqslant 2.
\end{aligned}$$

Then $\hat{\delta}_n' = \sup_{m\geqslant 1}\min_{1\leqslant i\leqslant m} f_{n+1}^{(m)}(i)\big/ f_n^{(m)}(i)$. For the first five of $\{\hat{\delta}_n'\}$, the minimum are all attained at m and so the computations become easier.

To compute $\bar{\hat{\delta}}_n$, simply use the formula

$$\bar{\hat{\delta}}_n = \sup_{m\geqslant 1} \frac{\sum_{i=1}^m 4^{1-i} f_n^{(m)}(i)^2 + 3^{-1} 4^{1-m} f_n^{(m)}(m)^2}{\sum_{i=1}^m 4^{2-i}\big(f_n^{(m)}(i) - f_n^{(m)}(i-1)\big)^2}, \qquad f_n^{(m)}(0) := 0. \quad \square$$

Example 5.4. *For Example 3.5 with* $\gamma = 1$ $\big(\hat{b}_i = i,\ \hat{a}_i = 2i\big)$, *we have* $\hat{\delta}_1' \approx 0.75 < \delta_1' \approx 0.84$ *and* $\hat{\delta}_1 \approx 1.12 > \delta_1 \approx 1.09$. *Besides,* $\hat{\delta}_1/\hat{\delta}_1' \approx 1.5$.

Example 5.5. *For Example 3.6, we have* $\hat{a}_i = \hat{b}_i = i^2\ (i \geqslant 1)$, $\hat{b}_0 = 0$, $\hat{\delta}_1' = 2 < \delta_1' \approx 2.19$ *and* $\hat{\delta}_1 = \delta_1 = 4$ *which is sharp. Besides,* $\hat{\delta}_1/\hat{\delta}_1' = 2$.

Proof. By Example 3.6 and Corollary 5.2, it follows that $\hat{\lambda}_0 = \lambda_0 = 1/4$. Here, we present an easier proof for the upper estimate. Note that when $\hat{a}_i = \hat{b}_i$ for $i \geqslant 2$, we have

$$\hat{\mu}_1 = 1,\ \ \hat{\mu}_i = \frac{\hat{b}_1 \cdots \hat{b}_{i-1}}{\hat{a}_2 \cdots \hat{a}_i} = \frac{\hat{b}_1}{\hat{a}_i},\ \ i \geqslant 2; \qquad \hat{\mu}_i \hat{b}_i = \hat{b}_1,\ \ i \geqslant 1. \tag{5.20}$$

In the present case, we have $\hat{\mu}_i = i^{-2}\ (i \geqslant 1)$ and $\hat{\mu}_i \hat{a}_i \equiv 1$. Let $f_i^{(m)} = \sqrt{i \wedge m}$. Then

$$\begin{aligned}
\hat{\mu}\big(f^{(m)\,2}\big) &= \sum_{i=1}^m \frac{1}{i} + m \sum_{i\geqslant m+1} \frac{1}{i^2},\\
\widehat{D}\big(f^{(m)}\big) &= \sum_{i=1}^m \big(\sqrt{i} - \sqrt{i-1}\,\big)^2 = \sum_{i=1}^m \frac{1}{\big(\sqrt{i} + \sqrt{i-1}\,\big)^2} \leqslant 1 + \frac{1}{4}\sum_{i=1}^{m-1}\frac{1}{i}.
\end{aligned}$$

Hence,

$$\hat{\lambda}_0 \leqslant \varliminf_{m\to\infty} \frac{\widehat{D}\big(f^{(m)}\big)}{\hat{\mu}\big(f^{(m)\,2}\big)} = \frac{1}{4}. \quad \square$$

Example 5.6. *For Example 3.7, we have* $\hat{a}_i = i^4\,(i \geqslant 1)$, $\hat{b}_i = i(i-1/2)(i^2+3i+3)$, $\hat{\lambda}_0 = \lambda_0 = 1/2$, $\hat{\delta}_1' \approx 1.83 < \delta_1' \approx 1.9$ *and* $\hat{\delta}_1 \approx \delta_1 \approx 2$. *Besides,* $\hat{\delta}_1/\hat{\delta}_1' = 1.09$.

Proof. First, we have

$$\hat{\mu}_i = \frac{\prod_{k=1}^{i-1}(k-1/2)(k^2+3k+3)}{i\,i!^3}, \quad \hat{\nu}_i = \frac{(i-1)!^3}{\prod_{k=1}^{i-1}(k-1/2)(k^2+3k+3)}, \quad i \geqslant 1.$$

By (5.5) and Example 3.7, we have $\sum_i \hat{\mu}_i < \infty$ and $\sum_i \hat{\nu}_i < \infty$, and so the minimal dual process is explosive (but here we are dealing with the maximal one). The sharp lower bound can be deduced from part (1) of Theorem 4.1 with the dual test sequence

$$\hat{v}_i = 1 + \frac{1}{i(i^2+3i+3)}, \qquad i \geqslant 1.$$

From this, it follows that the corresponding eigenfunction

$$\hat{g}_i = \prod_{k=1}^{i-1} \hat{v}_k, \quad i \geqslant 2, \quad \hat{g}_1 = 1,$$

increases strictly to a finite limit since $\sum_{i\geqslant 1} i^{-1}(i^2+3i+3)^{-1} < \infty$. The sequence $(\hat{v}_i)$ comes from the one computed in Example 3.7 plus a use of (2.35) and (5.7). $\square$

The precise value of λ_0 for the next example is unknown. Its eigenfunction is non-polynomial. It is interesting to compare this example with the ergodic one given in §6 for which $\lambda_1 = 2$, as well as the one with rates $a_i = i+1$ and $b_i = i^2\,(i \geqslant 1)$ given in §7 for which $\lambda_0 = 2$.

Example 5.7. *Let* $\hat{b}_0 = 0$, $\hat{b}_i = i+2\,(i \geqslant 1)$ *and* $\hat{a}_i = i^2$. *It is the dual of the process studied in §2 with rates* $a_i = i+2\,(i \geqslant 1)$ *and* $b_i = (i+1)^2\,(i \geqslant 0)$. *Then* $\hat{\lambda}_0 \in (0.395, 0.399)$, $\hat{\delta}_1' \approx 2.37 < \delta_1' \approx 2.48$ *and* $\hat{\delta}_1 \approx 2.63 > \delta_1 \approx 2.61$. *Besides,* $\hat{\delta}_1/\hat{\delta}_1' \approx 1.1$.

It is interesting that for all of Examples 5.3–5.7, we have $\hat{\delta}_1' \leqslant \delta_1'$ and $\hat{\delta}_1 \geqslant \delta_1$ which then means that Corollary 3.3 is more effective than Corollary 4.4. The effectiveness of the bounds δ_1 and δ_1' given in Corollary 4.4 was also checked by Sirl, Zhang and Pollett (2007) for some models from practice.

Remark 5.8. It is now a suitable position to mention a method for the numerical computation of λ_0 defined in §4. The idea is meaningful in the other cases. From proof (b) of Theorem 4.1, it follows that there is a sequence $(v_i : v_i > 1, 1 \leqslant i < N)$ such that

$$R_i(v) = a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i) = \lambda_0, \qquad v_0 = \infty,\ v_N = 0 \text{ if } N < \infty.$$

Hence, we have

$$\begin{cases} v_1 - 1 = (a_1 - \lambda_0) b_1^{-1}, \\ v_i - 1 = \big[a_i(1 - v_{i-1}^{-1}) - \lambda_0\big] b_i^{-1}, \quad 2 \leqslant i < N. \end{cases} \tag{5.21}$$

In other words, replacing $v_i - 1$ by u_i, when $z = \lambda_0$, the equation

$$\begin{cases} u_1 = (a_1 - z)b_1^{-1}, \\ u_i = \big[a_i u_{i-1}(1+u_{i-1})^{-1} - z\big]b_i^{-1}, \quad 2 \leqslant i < N, \end{cases} \tag{5.22}$$

has a positive solution $(u_i = u_i(z))_{1\leqslant i<N}$. Thus, one may use the maximal z so that (5.22) has a positive solution as an approximation of λ_0 (based on part (1) of Theorem 4.1). In this way, we obtain the approximation of $\hat{\lambda}_0$ given in Example 5.7.

6. Reflecting (Neumann) boundaries at origin and infinity (ergodic case)

We now turn to studying the first non-trivial eigenvalue in the ergodic case. Let $E = \{i : 0 \leqslant i < N+1\}$ ($N \leqslant \infty$), $b_0 > 0$, $b_N = 0$ if $N < \infty$,

$$\lambda_1 = \inf\big\{D(f) : \mu(f) = 0,\ \mu(f^2) = 1\big\}, \tag{6.1}$$

where $\mu(f) = \int f \mathrm{d}\mu$,

$$D(f) = \sum_{0\leqslant i<N} \mu_i b_i (f_{i+1} - f_i)^2 = \sum_{1\leqslant i<N+1} \mu_i a_i (f_i - f_{i-1})^2 \tag{6.2}$$

with domain $\mathscr{D}^{\max}(D) = \{f \in L^2(\mu) : D(f) < \infty\}$. In (6.1), we presume that

$$\sum_{i=0}^{N} \mu_i < \infty. \tag{6.3}$$

Then the Dirichlet form $(D, \mathscr{D}^{\max}(D))$ has a trivial eigenvalue $\lambda_0 = 0$ with constant eigenfunction $\mathbb{1}$, and here we are working on the next "eigenvalue" λ_1 of $(D, \mathscr{D}^{\max}(D))$. If (6.3) does not hold, then $\mathbb{1} \notin L^2(\mu)$ and so λ_1 is not meaningful. Moreover, by (1.3) and Proposition 1.3, the Dirichlet form is unique. In this case, the corresponding process is explosive, or zero-recurrent, or transient. The decay rate is described by λ_0 which has already been treated in Sections 2 and 3. Hence, throughout this section, we assume (6.3).

Note that condition (6.3) plus (1.2) means that the unique process is ergodic. When $N = \infty$ and (1.2) fails, the minimal process was treated in Sections 2 and 3, and in this section, we are dealing with the maximal process (cf. [10; Proposition 6.56]) as in Section 4, it is indeed the unique honest reversible process. Denote by $Q = (q_{ij})$ the birth–death Q-matrix. Then under (6.3), the maximal process $P_{ij}^{\max}(\lambda)$ (Laplace transform) can be expressed as

$$P_{ij}^{\max}(\lambda) = P_{ij}^{\min}(\lambda) + \frac{z_i(\lambda)\,\mu_j\,z_j(\lambda)}{\lambda\sum_k \mu_k\,z_k(\lambda)}, \qquad i, j \in E, \quad \lambda > 0,$$

where for each fixed j, $\{P_{ij}^{\min}(\lambda) : i \in E\}$ is the minimal solution to the equations

$$x_i = \sum_{k\neq i} \frac{q_{ik}}{\lambda + q_i} x_k + \frac{\delta_{ij}}{\lambda + q_i}, \qquad i \in E,$$

and $(z_i(\lambda) : i \in E)$ is the maximal solution to the equation

$$\begin{cases} (\lambda I - Q)u = 0, \\ 0 \leqslant u \leqslant 1, \end{cases} \qquad \lambda > 0$$

(cf. [10; Proposition 6.56]). According to a result due to Z.K. Wang (1964) (cf. Wang and Yang (1992, §6.8, Theorem 2)): if $N = \infty$ and (1.2) fails, then every honest process (may be non-symmetric) is ergodic and so is the maximal one. Certainly, within the symmetric context, by using (1.4), it is easy to check directly the ergodicity of the maximal process.

Here, we mention a technical point. If (6.3) fails, then as mentioned before, by (1.3), there is precisely one symmetrizable process (Dirichlet form) which is nothing but the minimal one. Thus, if (1.2) also fails, then the unique process must be explosive and so there is no honest symmetrizable process. This is a different point to the reversible case (i.e., (6.3) holds) for which there exists exactly one honest reversible process as just mentioned above.

We use the same notation I, II, $\mathscr{F}_I$, $\mathscr{F}_{II}$, $\widetilde{\mathscr{F}}_I$ and $\widetilde{\mathscr{F}}_{II}$ defined in Section 4 with an addition "$f_0 = 0$" in the last four sets, but redefine R and $\mathscr{V}$ as follows:

$$\begin{aligned} R_i(v) &= a_{i+1} + b_i - a_i/v_{i-1} - b_{i+1}v_i, \quad 0 \leqslant i < N, \\ &\qquad\qquad v_{-1} > 0 \text{ is free and so is } v_{N-1} \text{ if } N < \infty, \\ \mathscr{V} &= \{v : v_i > 0 \text{ for all } i : 0 \leqslant i < N-1\}. \end{aligned}$$

The local operator $\widetilde{R}$ is modified from R, replacing a_m by $\tilde{a}_m := \mu_m a_m \big/ \sum_{k=m}^N \mu_k$ for v with $\operatorname{supp}(v) = \{0, 1, \ldots, m-2\}$ in the set

$$\widetilde{\mathscr{V}} = \bigcup_{m=2}^{N} \bigg\{ v \colon \frac{a_{i+1}}{a_{i+2} + b_{i+1}} < v_i < \frac{a_{i+1} + b_i - a_i/v_{i-1}}{b_{i+1}} \ \text{ for } i = 0, 1, \ldots, m-2 \\ \text{and } v_i = 0 \text{ for } i \geqslant m-1 \bigg\}.$$

Theorem 6.1. *Under (6.3), the following variational formulas for λ_1 hold.*

(1) *Difference form:*

$$\inf_{v \in \widetilde{\mathscr{V}}} \sup_{0 \leqslant i < N} \widetilde{R}_i(v) = \lambda_1 = \sup_{v \in \mathscr{V}} \inf_{0 \leqslant i < N} R_i(v).$$

(2) *Summation form:*

$$\inf_{f \in \widetilde{\mathscr{F}}_I \cup \widetilde{\mathscr{F}}_I'} \sup_{1 \leqslant i \in E} I_i(\bar{f})^{-1} = \lambda_1 = \sup_{f \in \mathscr{F}_I} \inf_{1 \leqslant i \in E} I_i(\bar{f})^{-1},$$

where

$$\begin{aligned} \widetilde{\mathscr{F}}_I' &= \{f \in L^2(\mu) : f_0 = 0, f \text{ is strictly increasing}\}, \\ \bar{f} &= f - \pi(f), \quad \pi = \mu/Z \quad \text{and} \quad Z = \textstyle\sum_{i \in E} \mu_i. \end{aligned}$$

The use of $\bar f$ in the last line is based on the property $\bar f = \overline{f+c}$ for every constant c and so we can fix f_0 to be 0.

Proof of Theorem 6.1. In the ergodic case under (1.2), the assertion

$$\inf_{f\in\widetilde{\mathscr F}_I\cup\widetilde{\mathscr F}_I'}\ \sup_{1\leqslant i\in E} I_i(\bar f)^{-1}\geqslant \lambda_1$$

was proved in [7; Theorem 2.3] (but in the case that $k=\infty$ in the original proof, one requires the L^2-integrability condition included in $\widetilde{\mathscr F}_I'$, as was pointed out in proof (c) of Theorem 4.1). The proof remains the same in the present general situation with an obvious modification when $N<\infty$. Next, in the ergodic case under (1.2), the following result

$$\lambda_1=\sup_{v\in\mathscr V}\ \inf_{0\leqslant i<N} R_i(v)=\sup_{f\in\mathscr F_I}\ \inf_{1\leqslant i\in E} I_i(\bar f)^{-1} \tag{6.4}$$

is just [3; Theorem 1.1]. In the present general situation, the proof for the second equality in (6.4) needs a slight change only (cf. [3; Lemma 2.1]). To prove the first equality in (6.4), we claim that

$$\begin{aligned}\lambda_1&=\inf\big\{D(f):\mu\big(\big|f-\pi(f)\big|^2\big)=1,\ f_i=f_{i\wedge m}\ \text{for some}\ m\in E,\ m\geqslant 1\big\}\\ &=:\tilde\lambda_1.\end{aligned} \tag{6.5}$$

To see this, first it is clear that $\tilde\lambda_1\geqslant\lambda_1$. Next, the proof of [4; Theorem 3.2] gives us

$$\lambda_1\geqslant\sup_{f\in\mathscr F_I}\ \inf_{1\leqslant i\in E} I_i(\bar f)^{-1},$$

and furthermore, the equality sign with λ_1 replaced by $\tilde\lambda_1$ holds. Once again, the key point for the last statement is to show that the eigenfunction of $\tilde\lambda_1$ is strictly increasing. For this, the original proof needs only a modification replacing λ_1 by $\tilde\lambda_1$ (as indicated in proof (b) of Theorem 4.1). Therefore, (6.4) holds in the present general situation.

Now, we need only to show that

- $\inf_{f\in\widetilde{\mathscr F}_I}\sup_{1\leqslant i\in E} I_i(\bar f)^{-1}\leqslant \inf_{v\in\widetilde{\mathscr V}}\sup_{0\leqslant i<N}\widetilde R_i(v)$, and
- $\inf_{v\in\widetilde{\mathscr V}}\sup_{0\leqslant i<N}\widetilde R_i(v)\leqslant\lambda_1$.

(a) Prove that $\inf_{f\in\widetilde{\mathscr F}_I}\sup_{1\leqslant i\in E} I_i(\bar f)^{-1}\leqslant \inf_{v\in\widetilde{\mathscr V}}\sup_{0\leqslant i<N}\widetilde R_i(v)$.

As before, write $\widetilde R(u)$ instead of $\widetilde R(v)$. Given u with $\operatorname{supp}(u)=\{0,1,\ldots,m-1\}$ so that $(v_i)\in\widetilde{\mathscr V}$, where $v_i=u_{i+1}/u_i>0$ for $i<m-1$ and $v_i=0$ for $i\geqslant m-1$, let

$$f_i=\begin{cases} a_iu_{i-1}-b_iu_i, & i<m,\\ \tilde a_m u_{m-1}, & i\geqslant m.\end{cases}$$

Since the constraint in $\widetilde{\mathscr V}$ is equivalent to $\min_{i\leqslant m-1}\widetilde R_i(v)>0$, it is easy to check that

$$(f_{i+1}-f_i)/u_i=\widetilde R_i(u)>0\ \text{for}\ i<m\ \text{and}\ f_i=f_{i\wedge m},$$

and so $f + b_0 u_0 \in \widetilde{\mathscr{F}}_I$. Moreover, since

$$\begin{aligned}\sum_{k=i}^{N} \mu_k f_k &= \sum_{k=i}^{m-1} \mu_k f_k + f_m \sum_{j=m}^{N} \mu_j \\ &= \sum_{k=i}^{m-1} \mu_k (a_k u_{k-1} - b_k u_k) + f_m \sum_{j=m}^{N} \mu_j \\ &= \mu_i a_i u_{i-1} - \mu_m a_m u_{m-1} + \tilde{a}_m u_{m-1} \sum_{j=m}^{N} \mu_j \\ &= \mu_i a_i u_{i-1}, \qquad i \leqslant m-1,\end{aligned}$$

we get

$$\mu(f) = \sum_{k=0}^{N} \mu_k f_k = \mu_0 a_0 u_{-1} = 0,$$

and so

$$\begin{aligned}&\sum_{k=i}^{N} \mu_k \bar{f}_k = \mu_i a_i u_{i-1}, \qquad i \leqslant m-1, \\ &\sum_{k=m}^{N} \mu_k \bar{f}_k = f_m \sum_{k=m}^{N} \mu_k = \tilde{a}_m u_{m-1} \sum_{k=m}^{N} \mu_k = \mu_m a_m u_{m-1}.\end{aligned}$$

It follows that

$$u_{i-1} = \frac{1}{\mu_i a_i} \sum_{k=i}^{N} \mu_k \bar{f}_k, \qquad i \leqslant m.$$

Hence,

$$\widetilde{R}_{i-1}(u) = \frac{f_i - f_{i-1}}{u_{i-1}} = I_i(\bar{f})^{-1}, \qquad 1 \leqslant i \leqslant m.$$

Therefore, we have

$$\max_{0\leqslant i<m} \widetilde{R}_i(u) = \max_{1\leqslant i\leqslant m} I_i(\bar{f})^{-1} \geqslant \inf_{f\in\widetilde{\mathscr{F}}_I,\, f_i=f_{i\wedge m}} \max_{1\leqslant i\leqslant m} I_i(\bar{f})^{-1} \geqslant \inf_{f\in\widetilde{\mathscr{F}}_I} \sup_{1\leqslant i\in E} I_i(\bar{f})^{-1};$$

and then

$$\inf_{v\in\widetilde{\mathscr{V}}} \sup_{0\leqslant i<N} \widetilde{R}_i(v) \geqslant \inf_{f\in\widetilde{\mathscr{F}}_I} \sup_{1\leqslant i\in E} I_i(\bar{f})^{-1}.$$

Here, we have used the fact that $\widetilde{R}_i(v) = -\infty$ for $i \geqslant m-1$ if $\operatorname{supp}(v) = \{0, 1, \ldots, m-2\}$ and in the last step, we have returned to the original notation $\widetilde{R}(v)$ instead of $\widetilde{R}(u)$.

(b) Prove that $\inf_{v\in\widetilde{\mathscr{V}}} \sup_{0\leqslant i<N} \widetilde{R}_i(v) \leqslant \lambda_1$.

Because of

$$\big\{\mu(f) = 0,\ \mu\big(f^2\big) = 1,\ f_i = f_{i\wedge m}\big\} \subset \big\{\mu(f) = 0,\ \mu\big(f^2\big) = 1,\ f_i = f_{i\wedge(m+1)}\big\},$$

by (6.5), it is clear that

$$\lambda_1^{(m)} := \inf\big\{D(f): \mu(f)=0,\ \mu\big(f^2\big)=1,\ f_i=f_{i\wedge m}\big\} \downarrow \lambda_1 \ \text{ as } m\uparrow N.$$

Actually, this is a special case of an approximation result given in [2; Theorem 4.2 and Corollary 4.3] or [10; Theorem 9.20 and Corollary 9.21]. Note that $\lambda_1^{(m)}$ is just the first non-trivial eigenvalue of the local Dirichlet form $\big(\widetilde D,\, \mathscr{D}\big(\widetilde D\big)\big)$ defined by (4.17) replacing the Dirichlet boundary at 0 by the Neumann one (having the state space $\{0,1,\ldots,m\}$), with Neumann (reflecting) boundary at m. Denote by g the first eigenfunction of $\lambda_1^{(m)}$ and extend it to the whole space by setting $g_i=g_{i\wedge m}$. Now, if we set $u_i=g_{i+1}-g_i$ for $i\in E$, then $u_i>0$ for $i\leqslant m-1$, $u_i=0$ for $i\geqslant m$, and furthermore,

$$\widetilde R_i(u)=\lambda_1^{(m)}>0 \qquad \text{for all } i\leqslant m-1.$$

Moreover, by the definition of g, we have

$$\begin{cases} b_iu_i-a_iu_{i-1}=-\lambda_1^{(m)}g_i, & i\leqslant m-1,\\ \tilde a_m u_{m-1}=\lambda_1^{(m)}g_m. \end{cases}$$

Making a difference of this with the one replacing i by $i+1$, we get $\widetilde R_i(u)=\lambda_1^{(m)}$ for all $i\leqslant m-1$ (From this, the reason should be clear why in the definition of $\widetilde{\mathscr V}$, we use "$v_i=0$ for $i\geqslant m-1$" rather than "$v_i=v_{i\wedge m}$"). Thus,

$$\begin{aligned}
\lambda_1^{(m)} &= \max_{0\leqslant i<m} \widetilde R_i(u)\\
&\geqslant \inf_{u:\ \mathrm{supp}\,(u)=\{0,1,\ldots,m-1\};\ (v_i=u_{i+1}/u_i)\in\widetilde{\mathscr V}}\ \max_{0\leqslant i<m} \widetilde R_i(u)\\
&\geqslant \inf_{u:\ \mathrm{supp}\,(u)=\{0,1,\ldots,n\}\ \text{for some}\ n\geqslant 0,\ n<N;\ (v_i=u_{i+1}/u_i)\in\widetilde{\mathscr V}}\ \sup_{0\leqslant i<N} \widetilde R_i(u)\\
&= \inf_{v\in\widetilde{\mathscr V}}\ \sup_{0\leqslant i<N} \widetilde R_i(v).
\end{aligned}$$

Here in the last step, we have returned to the original notation $\widetilde R(v)$ instead of $\widetilde R(u)$. Letting $m\to N$, we obtain the required assertion. □

With the same rates (a_i,b_i) here but endow with the Dirichlet boundary at 0, we return to the situation studied in Section 4. The next result, taken from [7; Theorem 2.2] and [6; Theorem 3.5], is a comparison of λ_1 with the quantities λ_0, δ, δ_1 and δ_1' given in Section 4. See also Corollary 6.6 below for an improvement.

Theorem 6.2 (Criterion and basic estimates). *Under (6.3), $\lambda_1>0$ iff $\delta<\infty$. More precisely, we have*

$$\frac{1}{4\,\delta}\leqslant\frac{1}{\delta_1}\leqslant\lambda_0\leqslant\lambda_1\leqslant\lambda_0 Z\leqslant\frac{Z}{\delta_1'}\leqslant\frac{Z}{\delta}. \tag{6.6}$$

The next two results are mainly taken from [7; Theorem 2.4] with an addition on the monotonicity of $\{\eta_n\}$ and $\{\eta_n'\}$.

Theorem 6.3 (Approximating procedure). *Let* (6.3) *hold and* $\delta < \infty$. *Write* $\varphi_0 = 0$, $\varphi_i = \sum_{0 \leqslant j \leqslant i-1} (\mu_j b_j)^{-1} =: \nu[0, i-1]$ $(1 \leqslant i \in E)$, $\bar{f} = f - \pi(f)$, $\pi = \mu/Z$, *and* $Z = \sum_{k \in E} \mu_k =: \mu[0, N]$.

(1) *Define* $f_1 = \sqrt{\varphi}$, $f_n = \bar{f}_{n-1} I\!I(\bar{f}_{n-1})$ *and* $\eta_n = \sup_{1 \leqslant i \in E} I_i(\bar{f}_n)$. *Then* η_n *is decreasing in* n *and* $\lambda_1 \geqslant \eta_n^{-1} \geqslant (4\delta)^{-1}$ *for all* $n \geqslant 1$.

(2) *For fixed* $m \in E : m \geqslant 1$, *define*

$$f_1^{(m)} = \varphi_{\cdot \wedge m}, \qquad f_n^{(m)} = \big[\bar{f}_{n-1}^{(m)} I\!I\big(\bar{f}_{n-1}^{(m)}\big)\big](\cdot \wedge m), \qquad n \geqslant 2,$$

and then define

$$\eta_n' = \sup_{1 \leqslant m \in E} \inf_{1 \leqslant i \in E} I_i(\bar{f}_n), \qquad \bar{\eta}_n = \sup_{1 \leqslant m \in E} \frac{\mu\big(\bar{f}_n^{(m)\,2}\big)}{D\big(f_n^{(m)}\big)}, \qquad n \geqslant 1.$$

Then η_n' *is increasing in* n *and* $\eta_n'^{\,-1} \geqslant \bar{\eta}_n^{-1} \geqslant \lambda_1$ *for all* $n \geqslant 1$.

The notation "$\bar{f}_{n-1} I\!I(\bar{f}_{n-1})$" used in the theorem may have 0/0 but it should not cost any confusion. Note that here we use the same (ν_j) as in (2.15). In other words, when $b_0 > 0$, we use (2.15). But for its dual, it is more convenient to use $\hat{\nu}_j = (\hat{\mu}_j \hat{a}_j)^{-1}$ as in Section 4 since $b_0 = 0$. This is consistent with the notation used in Section 5.

As a consequence of Theorem 6.3, we have the following improvement of Theorem 6.2.

Corollary 6.4 (Improved estimates). *Let* (6.3) *hold. Then we have*

$$(4\delta)^{-1} \leqslant \eta_1^{-1} \leqslant \lambda_1 \leqslant \bar{\eta}_1^{-1}, \tag{6.7}$$

where

$$\eta_1 = \sup_{1 \leqslant i \in E} \big(\sqrt{\varphi_i} + \sqrt{\varphi_{i-1}}\big)\left[\psi_i - \psi_1 \frac{\mu[i, N]}{\mu[0, N]}\right], \quad \psi_i := \sum_{j=i}^{N} \mu_j \sqrt{\varphi_j}, \tag{6.8}$$

$$\begin{aligned} \bar{\eta}_1 &= \sup_{1 \leqslant m \in E} \frac{1}{\varphi_m}\left[\sum_{1 \leqslant k \in E} \mu_k \varphi_{k \wedge m}^2 - \frac{1}{Z}\left(\sum_{1 \leqslant k \in E} \mu_k \varphi_{k \wedge m}\right)^2\right] \\ &= \sup_{1 \leqslant m \in E} \left\{\frac{1}{\varphi_m}\left[\sum_{1 \leqslant k \leqslant m-1} \mu_k \varphi_k^2 - \frac{1}{\mu[0, N]}\left(\sum_{1 \leqslant k \leqslant m-1} \mu_k \varphi_k\right)^2\right] \right. \\ &\qquad\qquad \left. + \frac{\mu[m, N]}{\mu[0, N]}\left[\varphi_m \mu[0, m-1] - 2 \sum_{1 \leqslant k \leqslant m-1} \mu_k \varphi_k\right]\right\}. \end{aligned} \tag{6.9}$$

Proof of Theorem 6.3.

Part 1. We prove that $\{f_n\} \subset L^1(\mu)$ in three steps. This was missed in the original paper [7]. Certainly, we need only to consider the case that $N = \infty$.

(a) First, we show that the functions $\{h_n\}$,

$$h_0(i) \equiv 1, \quad i \in E, \qquad h_n(i) = \sum_{j=1}^{i} \frac{1}{\mu_j a_j} \sum_{k=j}^{\infty} \mu_k h_{n-1}(k), \quad i \geqslant 1,\ n \geqslant 1,$$

are all in $L^1(\mu)$. Clearly, h_1 (and then h_n for $n \geqslant 2$) may increase to infinity if the minimal process is recurrent which is the main problem we need to handle. The required assertion says that even though h_n can be unbounded but is still in $L^1(\mu)$. For this, to distinguish with $\{f_n\}$ used in Theorem 6.3, let $\{\tilde{f}_n\}$ be the sequence defined in part (1) of Theorem 4.3:

$$\tilde{f}_1(i) = \left(\sum_{k=1}^{i} \nu_{k-1}\right)^{1/2} = f_1(i), \qquad i \geqslant 1,\ \nu_{j-1} := \frac{1}{\mu_j a_j},$$
$$\tilde{f}_n(i) = \sum_{j=1}^{i} \nu_{j-1} \sum_{k=j}^{\infty} \mu_k \tilde{f}_{n-1}(k), \qquad i \geqslant 1,\ n \geqslant 2,$$
$$\tilde{f}_n(0) = 0, \qquad n \geqslant 1.$$

From proof (b) of Theorem 4.3, we have seen that

$$\tilde{f}_2(i) = \sum_{j=1}^{i} \frac{1}{\mu_j a_j} \sum_{k \geqslant j} \mu_k \tilde{f}_1(k) \leqslant 4\delta \tilde{f}_1(i).$$

Because $\tilde{f}_1(i) \geqslant \tilde{f}_1(1) = a_1^{-1/2}$ for $i \geqslant 1$, this gives us

$$h_1(i) \leqslant 4\delta \sqrt{a_1}\, \tilde{f}_1(i), \qquad i \geqslant 1.$$

By induction, it follows that

$$h_n \leqslant \sqrt{a_1}\, (4\delta)^n \tilde{f}_1, \qquad n \geqslant 1.$$

This proves that $h_n \in L^1(\mu)$ for all $n \geqslant 1$ since $\tilde{f}_1 \in L^1(\mu)$ as mentioned in proof (b) of Theorem 4.3, due to the assumption $\delta < \infty$.

(b) Next, we study the relation between $\{f_n\}$ and $\{\tilde{f}_n\}$. By definition, we have

$$f_2(i) = \sum_{j=1}^{i} \nu_{j-1} \sum_{k=j}^{\infty} \mu_k \bar{f}_1(k) = \tilde{f}_2(i) - h_1(i)\pi(f_1), \qquad i \geqslant 1,$$
$$f_3(i) = \sum_{j=1}^{i} \nu_{j-1} \sum_{k=j}^{\infty} \mu_k \bar{f}_2(k) = \tilde{f}_3(i) - h_2(i)\pi(f_1) - h_1(i)\pi(f_2), \qquad i \geqslant 1,$$
$$f_4(i) = \tilde{f}_4(i) - h_3(i)\pi(f_1) - h_2(i)\pi(f_2) - h_1(i)\pi(f_3), \qquad i \geqslant 1.$$

Successively, we obtain

$$f_n = \tilde{f}_n - \sum_{k=1}^{n-1} \pi(f_k) h_{n-k}, \qquad n \geqslant 2.$$

(c) Since $f_1 = \tilde{f}_1 \in L^1(\mu)$ as shown in proof (b) of Theorem 4.3. Now, to show that $\{f_n\} \subset L^1(\mu)$, by (a) and (b), it suffices to prove that $\{\tilde{f}_n\} \subset L^1(\mu)$. This is done in proof (b) of Theorem 4.3.

Part 2. We now prove the monotonicity of $\{\eta_n\}$ in two steps. Since $\bar{f}_n$ values both positive and negative or even zero, the proportional property used in the proof of the monotonicity of $\{\delta_n\}$ is currently not available. To overcome this difficulty, a finer technique is needed.

(d) Because

$$\mu_i a_i \big[f_n(i) - f_n(i-1)\big] = \sum_{k=i}^{N} \mu_k \bar{f}_{n-1}(k), \qquad n \geqslant 2,$$

by the definition of $I(f)$, we obtain

$$\eta_n = \sup_{1 \leqslant i \in E} \sum_{j=i}^{N} \mu_j \bar{f}_n(j) \bigg/ \sum_{k=i}^{N} \mu_k \bar{f}_{n-1}(k), \qquad n \geqslant 2. \tag{6.10}$$

Since the denominator is positive, the assertion that $\eta_n \leqslant \eta_{n-1}$ is equivalent to

$$\sum_{j=i}^{N} \mu_j \big[\bar{f}_n(j) - \eta_{n-1} \bar{f}_{n-1}(j)\big] \leqslant 0, \qquad i \in E.$$

That is,

$$\eta_{n-1}\pi(f_{n-1}) - \pi(f_n) \leqslant \frac{1}{\mu[i,N]} \sum_{j=i}^{N} \mu_j \big[\eta_{n-1} f_{n-1}(j) - f_n(j)\big], \quad i \in E. \tag{6.11}$$

Let us observe the meaning of this inequality: the left-hand side is the infimum (attained at $i = 0$) of the right-hand side.

The monotonicity of $\{\eta_n\}$ now follows once we show that the right-hand side of (6.11) is luckily increasing in i, or equivalently,

$$\mu[i,N] \sum_{j=i+1}^{N} \mu_j \big[\eta_{n-1} f_{n-1}(j) - f_n(j)\big] \geqslant \mu[i+1,N] \sum_{j=i}^{N} \mu_j \big[\eta_{n-1} f_{n-1}(j) - f_n(j)\big].$$

By removing the common term

$$\mu[i+1,N] \sum_{j=i+1}^{N} \mu_j \big[\eta_{n-1} f_{n-1}(j) - f_n(j)\big]$$

in both sides, it is enough to check that

$$\eta_{n-1} \sum_{j=i+1}^{N} \mu_j \big[f_{n-1}(j) - f_{n-1}(i)\big] \geqslant \sum_{j=i+1}^{N} \mu_j \big[f_n(j) - f_n(i)\big], \quad i \in E,\ n \geqslant 2. \tag{6.12}$$

First, let $n \geqslant 3$. Then by the definition of f_n and (6.10), we have

$$\begin{aligned} f_n(j) - f_n(i) &= \sum_{s=i+1}^{j} \nu_{s-1} \sum_{k=s}^{N} \mu_k \bar{f}_{n-1}(k) \\ &\leqslant \eta_{n-1} \sum_{s=i+1}^{j} \nu_{s-1} \sum_{k=s}^{N} \mu_k \bar{f}_{n-2}(k) \\ &= \eta_{n-1} \big[f_{n-1}(j) - f_{n-1}(i)\big]. \end{aligned}$$

This certainly implies (6.12) in the case of $n \geqslant 3$, regarded as an application of the proportional property. Next, let $n = 2$. Then by the definition of f_2 and η_1, we have

$$\begin{aligned} f_2(j) - f_2(i) &= \sum_{s=i+1}^{j} \nu_{s-1} \sum_{k=s}^{N} \mu_k \bar{f}_1(k) \\ &\leqslant \eta_1 \sum_{s=i+1}^{j} \big[f_1(s) - f_1(s-1)\big] \\ &= \eta_1 \big[f_1(j) - f_1(i)\big]. \end{aligned}$$

This also implies (6.12) in the case of $n = 2$. We have thus proved that $\eta_n \leqslant \eta_{n-1}$ for all $n \geqslant 2$.

Part 3. To prove the monotonicity of $\{\eta_n'\}$, for each fixed m, as a dual argument (exchanging "sup" and "$\leqslant$" with "inf" and "$\geqslant$", respectively) of the above proofs (d) and (e), we have

$$\inf_{1 \leqslant i \in E} I_i\big(\bar{f}_n^{(m)}\big) \leqslant \inf_{1 \leqslant i \in E} I_i\big(\bar{f}_{n+1}^{(m)}\big).$$

Then the assertion follows by making supremum with respect to m.

Part 4. The proof of $\bar{\eta}_n \geqslant \eta_n'$ is given in Lemma 6.5 below. □

In practice, using $\bar{\eta}_n$ rather than η_n' is based on the following result.

Lemma 6.5. *For every non-decreasing, and non-constant function f satisfying $f \in L^1(\mu)$ and $D(f) < \infty$, we have*

$$\frac{\mu\big(\bar{f}^2\big)}{D(f)} \geqslant \inf_{1 \leqslant i \in E} I_i(\bar{f}).$$

Similarly, for every nonnegative, non-decreasing, and non-zero function f satisfying $f \in L^1(\mu)$ and $D(f) < \infty$, we have

$$\frac{\mu\big(f^2\big)}{D(f)} \geqslant \inf_{1 \leqslant i \in E} I_i(f).$$

Proof. (a) Since f is not a constant, we have $\mu(\bar f^2)>0$ and $D(f)>0$. Moreover, since $f\in L^1(\mu)$ is also non-decreasing, we claim that

$$\infty>\sum_{k=i}^N \mu_k\bar f_k>0\qquad\text{for all } i\in E,\qquad i\geqslant 1.$$

Actually, the non-decreasing sequence $\{\bar f_k\}$, starting at $\bar f_0<0$ (since f is non-trivial) and having mean zero, should be positive for all large enough k. Thus, if $\sum_{k=i_0}^N\mu_k\bar f_k\leqslant 0$ for some $i_0: 1\leqslant i_0\in E$, then we would have $\bar f_{i_0}<0$ $\big($otherwise $\bar f_i\geqslant 0$ for all $i\geqslant i_0$ and then $\sum_{k=i_0}^N\mu_k\bar f_k>\mu_j\bar f_j>0$ for large enough $j\big)$. This implies that

$$\sum_{k\leqslant i_0-1}\mu_k\bar f_k\leqslant \bar f_{i_0}\sum_{k\leqslant i_0-1}\mu_k<0,$$

and furthermore,

$$0=\mu(\bar f)=\sum_{k=i_0}^N\mu_k\bar f_k+\sum_{k\leqslant i_0-1}\mu_k\bar f_k\leqslant\sum_{k\leqslant i_0-1}\mu_k\bar f_k<0,$$

which is a contradiction. Because of the assertion we have just proved and using the convention that $1/0=\infty$, it follows that $\inf_{1\leqslant i\in E}I_i(\bar f)\in[0,\infty)$.

Let $\gamma=\inf_{1\leqslant i\in E}I_i(\bar f)$. Then we have

$$-\sum_{k\leqslant i-1}\mu_k\bar f_k=\sum_{k=i}^N\mu_k\bar f_k\geqslant\gamma\,\mu_i a_i(\bar f_i-\bar f_{i-1})$$

first for those i with $f_i>f_{i-1}$ and then for all $i: 1\leqslant i\in E$. Multiplying both sides by $\bar f_i-\bar f_{i-1}\geqslant 0$, we obtain

$$-(\bar f_i-\bar f_{i-1})\sum_{k\leqslant i-1}\mu_k\bar f_k\geqslant\gamma\,\mu_i a_i(\bar f_i-\bar f_{i-1})^2,\qquad i\in E,\ i\geqslant 1.$$

Making a summation over i from 1 to m, it follows that

$$-\sum_{i=1}^m(\bar f_i-\bar f_{i-1})\sum_{k\leqslant i-1}\mu_k\bar f_k\geqslant\gamma\sum_{i=1}^m\mu_i a_i(\bar f_i-\bar f_{i-1})^2.$$

Noticing that the mean of $\bar f$ equals zero and exchanging the order of the sums, the left-hand side is equal to

$$\begin{aligned}-\sum_{k=0}^m\mu_k\bar f_k\sum_{i=k+1}^m(\bar f_i-\bar f_{i-1})&=-\sum_{k=0}^m\mu_k\bar f_k(\bar f_m-\bar f_k)\\&=-\bar f_m\sum_{k=0}^m\mu_k\bar f_k+\sum_{k=0}^m\mu_k\bar f_k^2\\&=\sum_{k=m+1}^N\mu_k\bar f_k\bar f_m+\sum_{k=0}^m\mu_k\bar f_k^2.\end{aligned}$$

As mentioned in the last paragraph, $\bar{f}_m > 0$ first for some m and then for all large enough m since $\bar{f}$ is non-decreasing, the right-hand side is controlled, for large enough m, from above by

$$\sum_{k=m+1}^{N} \mu_k \bar{f}_k^2 + \sum_{k=0}^{m} \mu_k \bar{f}_k^2 = \mu(\bar{f}^2).$$

With the assumption $D(f) < \infty$ in mind, the required assertion now follows immediately by passing the limit as $m \to N$.

(b) For the second assertion, since $f \in L^1(\mu)$ is nonnegative and non-zero, we have

$$\infty > \sum_{j=i}^{N} \mu_j f_j > 0 \qquad \text{for all } i \in E.$$

Now, if $f_0 = 0$, then there is an i_0 such that $f_{i_0-1} = 0$ but $f_{i_0} > 0$ and so $I_{i_0}(f) < \infty$. If $f_0 > 0$ and $\inf_{1 \leqslant i \in E} I_i(f) = \infty$, then f should be a positive constant, and hence, $D(f) = 0$. In this case, the assertion is trivial since $\mu(f^2) > 0$. Therefore, we may assume that $\gamma := \inf_{1 \leqslant i \in E} I_i(f) < \infty$. We now have

$$\sum_{k=i}^{N} \mu_k f_k \geqslant \gamma \mu_i a_i (f_i - f_{i-1}), \qquad i \in E,\ i \geqslant 1.$$

Hence,

$$\sum_{i=1}^{m} (f_i - f_{i-1}) \sum_{k=i}^{N} \mu_k f_k \geqslant \gamma \sum_{i=1}^{m} \mu_i a_i (f_i - f_{i-1})^2.$$

Exchanging the order of the sums, the left-hand side is equal to

$$\sum_{k=1}^{N} \mu_k f_k \sum_{i=1}^{k\wedge m} (f_i - f_{i-1}) = \sum_{k=1}^{N} \mu_k f_k (f_{k\wedge m} - f_0) \leqslant \sum_{k=1}^{N} \mu_k f_k f_{k\wedge m} \leqslant \mu(f^2).$$

Combining this with the last inequality, we have obtained the required assertion. $\square$

Having the comparison of $\bar{\eta}_n \geqslant \eta_n'$ (Lemma 6.5) in mind, one may expect a parallel result for δ_n' and $\bar{\delta}_n$ defined in Theorem 4.3. All the examples we have ever computed support the conjecture that $\bar{\delta}_n \geqslant \delta_n'$, however, there is still no proof. In general, we have $\bar{\delta}_{n+1} \geqslant \delta_n'$ only as stated in Theorem 4.3. Note that δ_n' is defined by using $I\!I(f_n)$ rather than $I(f_n)$. If we redefine δ_n' by using $I(f_n)$ as in [7; Theorem 2.2], denoted by $\tilde{\delta}_n'$ for a moment, then by the second assertion of Lemma 6.5, we do have $\bar{\delta}_n \geqslant \tilde{\delta}_n'$. Besides, by the theorem just quoted, we also have $\delta_n' \geqslant \tilde{\delta}_n' \geqslant \delta_{n-1}'$. This remark is also meaningful for those δ_n' and $\bar{\delta}_n$ defined in Section 3.

Note that the factor of the upper and lower bounds of λ_1 given in Theorem 6.2 is $4Z > 4$. The next result has a factor 4 only. A simple comparison of κ below and $\delta^{(4.4)}$ shows that it is not easy to find such a result. Its proof is delayed to the next section.

Corollary 6.6 (Criterion and basic estimates). *Let* (6.3) *hold. Then we have* $\kappa^{-1}/4 \leqslant \lambda_1 \leqslant \kappa^{-1}$, *where*

$$\kappa^{-1} = \inf_{0\leqslant n<m<N+1} \left[\left(\sum_{i=0}^{n} \mu_i\right)^{-1} + \left(\sum_{i=m}^{N} \mu_i\right)^{-1}\right]\left(\sum_{j=n}^{m-1} \frac{1}{\mu_j b_j}\right)^{-1}. \tag{6.13}$$

Furthermore, we have

$$\delta_L \wedge \delta_R \geqslant \kappa \geqslant Z^{-1}\delta_L,$$

where $Z = \sum_{i=0}^N \mu_i$,

$$\delta_L = \sup_{1\leqslant n<N+1} \sum_{i=1}^{n} \frac{1}{\mu_i a_i} \sum_{j=n}^{N} \mu_j = \delta^{(4.4)}, \qquad \delta_R = \sup_{0\leqslant m<N} \sum_{j=0}^{m} \mu_j \sum_{k=m}^{N-1} \frac{1}{\mu_k b_k}.$$

In the case that the minimal process is ergodic, since

$$1 < Z < \infty, \qquad \sum_j \frac{1}{\mu_j b_j} = \infty,$$

we have $\delta_R = \infty$ and so the second assertion of Corollary 6.6 goes back to Theorem 6.2. However, the first assertion of Corollary 6.6 is clearly finer. An extension of Corollary 6.6 to a more general state space is given in Corollary 7.9 below.

Most of the examples below are taken from [10; Examples 9.27]. The computation of $\bar\eta_1$, $\eta_1/\bar\eta_1$, and κ is newly added.

Example 6.7. *Let* $b_i = b\,(i \geqslant 0)$, *and* $a_i = a\,(i \geqslant 1)$, $a > b$. *Then*

$$\lambda_1 = \big(\sqrt{a} - \sqrt{b}\,\big)^2, \quad \delta = \kappa = a(a-b)^{-2}, \quad \bar\eta_1 = \delta_1' = (a+b)/(a-b)^2,$$

and $\eta_1 = \lambda_1^{-1}$ *which is sharp. Besides,* $\eta_1/\bar\eta_1 \leqslant 2$, *the equality sign holds iff* $b = a$. *Note that* λ_1, η_1^{-1}, *and* $\bar\eta_1^{-1}$ *all tend to zero as* $b \to a$. *Furthermore,* $(\bar\eta_1, \eta_1) \subset (\kappa, 4\kappa)$.

Example 6.8. *The typical linear model: let* $b_i = \beta_1 i + \beta_0\,(\beta_0 > 0,\ \beta_1 \geqslant 0)$, *and* $a_i = \gamma_1 i\,(\gamma_1 > \beta_1)$ *for* $i \geqslant 0$. *Then* $\lambda_1 = \gamma_1 - \beta_1$. *When* $\beta_0 = 0$, *we have* $\lambda_0^{(4.2)} = \gamma_1 - \beta_1$.

Example 6.9. *Let* $b_i = b/(i+1)\,(b > 0)$ *for* $i \geqslant 0$, $a_i \equiv a > 0$ *for* $i \geqslant 1$. *Then* $\lambda_1 = a - \big(\sqrt{b^2 + 4ab} - b\big)/2$.

Example 6.10. *Let* $b_i \equiv b\,(b > 0)$ *for* $i \geqslant 0$, $a_i = (i \wedge k)a\,(a > 0)$ *for* $i \geqslant 1$ *and some* $k \geqslant 2$ *satisfying*

$$k^{-1} \leqslant a/b \leqslant k(k-1)^{-2}.$$

Then $\lambda_1 = \big(\sqrt{ak} - \sqrt{b}\big)^2$.

Example 6.11. *Let $b_0 = 1$, $b_i = i$, and $a_i = 2\,i$, $i \geqslant 1$. Then $\lambda_1 \geqslant \lambda_0 = 1$ but the precise value is unknown. Moreover,*

$$\bar{\eta}_1 \approx 0.55, \quad \eta_1 \approx 0.9986 \quad \text{and} \quad \eta_1/\bar{\eta}_1 \approx 1.82 < 2.$$

Besides, $\kappa \approx 0.4856$ and so $(\bar{\eta}_1, \eta_1) \subset (\kappa, 4\kappa)$.

The next one is a continuation of [6; Example 3.10].

Example 6.12. *Let $E = \{0, 1\}$. Then $\lambda_1 = Z\lambda_0 = \bar{\eta}_1^{-1} = \kappa^{-1}$ and $\lambda_0 = \delta^{-1}$. Hence, the last upper bound in (6.6) and the one in Corollary 6.6 are sharp but δ^{-1} is not an upper bound of λ_1.*

The first lower bound in (6.6) and the one in Corollary 6.6 are sharp for the seventh example in Table 6.1 below.

Examples 6.13. *Here are some additional examples, given in Table 6.1, for which the quantities $\bar{\eta}_1 \leqslant \lambda_1^{-1} \leqslant \eta_1$ and $\kappa \leqslant \lambda_1^{-1} \leqslant 4\kappa$ are compared. For all these examples, we have $(\bar{\eta}_1, \eta_1) \subset (\kappa, 4\kappa)$ and so the estimates given in Corollary 6.4 are better than the ones in Corollary 6.6.*

Table 6.1 Exact λ_1 and its estimates for eight examples

b_i $(i \geqslant 0)$	a_i $(i \geqslant 1)$	λ_1^{-1}	$\bar{\eta}_1$	η_1	$\eta_1/\bar{\eta}_1$	κ
$i+1$	$2i$	1	≈ 0.8	≈ 1.48	≈ 1.85	$2/3$
$i+1$	$2i+3$ $a_0 = 0$	$1/2$	≈ 0.346	≈ 0.638	≈ 1.84	≈ 0.28
$i+1$	$2i+4+\sqrt{2}$ $a_0 = 0$	$1/3$	≈ 0.218	≈ 0.398	≈ 1.83	≈ 0.18
$(i+1)^{-1}$	1	$2(3-\sqrt{5})^{-1}$ ≈ 2.618	≈ 1.92	≈ 3.24	≈ 1.69	≈ 1.6
1	$i \wedge 2$	$(\sqrt{2}-1)^{-2}$ ≈ 5.8284	≈ 3	≈ 5.8284	≈ 1.9	2
$i+2$	i^2	$1/2$	≈ 0.47	≈ 0.85	≈ 1.81	≈ 0.47
i^2 $b_0 = 1$	i^2	4	2	λ_1^{-1}	2	1
$2+(-1)^i$ $b_0 = \dfrac{7-\sqrt{33}}{2}$	$2[2+(-1)^i]$ $a_0 = 0$	$(6-\sqrt{33})^{-1}$ ≈ 3.9	≈ 2.11	≈ 4.21	≈ 2	≈ 1.56

7. Bilateral absorbing (Dirichlet) boundaries

This section deals with the fourth case of boundary conditions. It consists of two parts. The first one is for the ordinary birth–death processes as studied in the previous sections and the second one deals with the bilateral birth–death processes with a more general state space.

First, let us consider the processes with state space $E = \{i : 1 \leqslant i < N + 1\}$ $(N \leqslant \infty)$ with Dirichlet boundaries at 0 $(a_1 > 0)$ and $N+1$ if $N < \infty$. Similar to Section 2, define

$$\lambda_0 = \inf\big\{D(f) : \mu(f^2) = 1,\ f \in \mathscr{K}\big\}, \tag{7.1}$$

where the symmetric measure (μ_i) is the same as in Section 4, $\mu(f) = \sum_{k\in E} \mu_k f_k$, and

$$D(f) = \sum_{k\in E} \mu_k a_k (f_k - f_{k-1})^2, \qquad f_0 := 0,$$

with domain $\mathscr{D}^{\min}(D)$. Clearly, if one changes only the boundary condition at 0, then the resulting λ_0 is bigger or equal to $\lambda_0^{(2.2)}$. Note that if (1.3) fails, then the eigenvalues $\lambda_0^{(4.2)}$ and $\lambda_0^{(7.1)}$ are different which correspond to the maximal and the minimal Dirichlet forms, respectively. However, as mentioned in Section 4, once (1.3) holds, $\lambda_0^{(4.2)}$ coincides with $\lambda_0^{(7.1)}$. Then there are three cases. The first one is that $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i a_i)^{-1} = \infty$. This case is treated in Section 4. In this section, we are mainly studying the second case that $\sum_i \mu_i = \infty$ but

$$\sum_{k=1}^{N} \frac{1}{\mu_k a_k} < \infty. \tag{7.2}$$

The third case is that $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i a_i)^{-1} = \infty$ which is treated in the next theorem. In this degenerated case, since there is a killing at 1 (i.e., $a_1 > 0$), the process is transient. Without using duality, by Corollary 7.3 below, we also obtain that $\lambda_0 = 0$. See the comments right after Corollary 7.3.

Theorem 7.1.

(1) *First, let* (7.2) *hold. Define the dual rates* $\big(\hat a_i, \hat b_i\big)$ *by* (5.1) *in the inverse way:*

$$\hat b_i = a_{i+1}, \qquad \hat a_i = b_i, \qquad 0 \leqslant i < N+1, \tag{7.3}$$

and denote by $\hat\lambda_1$ *the eigenvalue defined in Section 6 for the dual process. Then we have* $\lambda_0 = \hat\lambda_1$.

(2) *Next, let (7.2) fail and* $\sum_{i\in E} \mu_i = \infty$. *Then* λ_0 $\big($*as well as* $\lambda_0^{(4.2)}\big)$ *is equal to its dual* $\hat\lambda_0 = \lambda_0^{(2.2)} = 0$.

Proof. (a) By (7.3), (7.2) and (5.5), we have

$$\sum_{k=0}^{N} \hat\mu_k < \infty. \tag{7.4}$$

Clearly, the dual process with rates $\big(\hat a_i, \hat b_i\big)$ has the state space $\widehat{E} = \{i : 0 \leqslant i < N+1\}$. By exchanging (a_i, b_i, v_i) and $\big(\hat a_i, \hat b_i, \hat v_i\big)$ in part (1) of Theorem 6.1,

$$\hat\lambda_1 = \sup_{\hat v} \inf_{0\leqslant i<N} \bigg[\hat a_{i+1} + \hat b_i - \frac{\hat a_i}{\hat v_{i-1}} - \hat b_{i+1}\hat v_i\bigg],$$

and in (5.8) with $N' = N$,

$$\sup_{\hat v}\inf_{0\leqslant i<N}\left[\hat a_{i+1}+\hat b_i-\frac{\hat a_i}{\hat v_{i-1}}-\hat b_{i+1}\hat v_i\right]=\sup_v\inf_{i\in E}\left[a_i\left(1-\frac{1}{v_{i-1}}\right)+b_i(1-v_i)\right],$$

the first assertion of Theorem 7.1 now follows from the variational formula given on the right-hand side of (9.2) in Section 9.

(b) Similarly, replacing the use of Theorem 6.1 by Proposition 2.7 (1), we obtain the second assertion. In this case, as already mentioned at the beginning of Section 4, we have $\lambda_0 = \lambda_0^{(4.2)}$. The fact that $\hat\lambda_0 = \lambda_0^{(2.2)} = 0$ comes from (5.5) and Theorem 3.1. $\square$

By Theorem 7.1 (1), all the results obtained in Section 6 can be transformed into the present setup. For instance, by Corollary 6.4, we obtain the following result.

Corollary 7.2. *Under (7.2), we have* $(4\delta)^{-1}\leqslant \delta_1^{-1}\leqslant\lambda_0\leqslant\bar\delta_1^{-1}$*, where*

$$\delta=\left[\sup_{1\leqslant n<N}\mu[1,n]\left(\nu[n+1,N]+\frac{\mathbb{1}_{\{N<\infty\}}}{\mu_N b_N}\right)\right]\bigvee\frac{\mu[1,N]\,\mathbb{1}_{\{N<\infty\}}}{\mu_N b_N},\qquad \nu_k=\frac{1}{\mu_k a_k},$$

$$\delta_1=\sup_{i\in E}\left(\sqrt{\varphi_i}+\sqrt{\varphi_{i-1}}\right)\left(\psi_i-\psi_1\frac{\nu[i+1,N]+(\mu_N b_N)^{-1}\mathbb{1}_{\{N<\infty\}}}{\nu[1,N]+(\mu_N b_N)^{-1}\mathbb{1}_{\{N<\infty\}}}\right),$$

$$\bar\delta_1=\sup_{m\in E}\frac{1}{\varphi_m}\Bigg[\sum_{k=1}^{N-1}\nu_{k+1}\varphi_{k\wedge m}^2+\frac{\varphi_m^2}{\mu_N b_N}\,\mathbb{1}_{\{N<\infty\}}-$$
$$-\frac{1}{\nu[1,N]+(\mu_N b_N)^{-1}\mathbb{1}_{\{N<\infty\}}}\left(\sum_{k=1}^{N-1}\nu_{k+1}\varphi_{k\wedge m}+\frac{\varphi_m}{\mu_N b_N}\,\mathbb{1}_{\{N<\infty\}}\right)^2\Bigg],$$

$$\varphi_i=\mu[1,i],\qquad \psi_i=\sum_{j=i}^{N-1}\nu_{j+1}\sqrt{\varphi_j}+\frac{1}{\mu_N b_N}\sqrt{\varphi_N}\,\mathbb{1}_{\{N<\infty\}},\qquad i\in E.$$

Proof. Starting from Corollary 6.4 with its notation, write everything we need in its dual. First by (5.4), we have

$$\mu_n=\hat a_1\hat\nu_{n+1},\quad \nu_n=\frac{1}{\hat a_1}\hat\mu_{n+1},\quad 0\leqslant n<N,\qquad \mu_N=\frac{\hat a_1}{\hat\mu_N\hat b_N}\quad\text{if } N<\infty.$$

Here, recall that $\nu_n=(\mu_n b_n)^{-1}$ but $\hat\nu_n=\left(\hat\mu_n\hat a_n\right)^{-1}$. Then the constant δ defined

in (4.4) becomes $\delta = \sup_{n\in E} \nu[0, n-1]\, \mu[n, N]$. Moreover, we have

$$\begin{aligned}
\varphi_i &= \sum_{j=0}^{i-1} \nu_j = \frac{1}{\hat a_1}\hat\mu[1,i], \qquad 1 \leqslant i < N+1,\\
\mu[m,n] &= \sum_{j=m}^{n} \mu_j = \hat a_1 \hat\nu[m+1, n+1], \qquad 0 \leqslant m \leqslant n < N,\\
\mu[m,N] &= \hat a_1 \hat\nu[m+1,N] + \mu_N \mathbb{1}_{\{N<\infty\}} = \hat a_1\biggl[\hat\nu[m+1,N] + \frac{\mathbb{1}_{\{N<\infty\}}}{\hat\mu_N \hat b_N}\biggr], \qquad m<N,\\
\psi_i &= \sum_{j=i}^{N-1} \mu_j \sqrt{\varphi_j} + \mu_N \sqrt{\varphi_N}\, \mathbb{1}_{\{N<\infty\}}\\
&= \sqrt{\hat a_1}\biggl[\sum_{j=i}^{N-1} \hat\nu_{j+1} \sqrt{\hat\mu[1,j]} + \frac{1}{\hat\mu_N \hat b_N}\sqrt{\hat\mu[1,N]}\, \mathbb{1}_{\{N<\infty\}}\biggr].
\end{aligned}$$

Inserting these quantities into (6.8) and (6.9), making a little simplification, and then ignoring the hat everywhere, we obtain Corollary 7.2. $\square$

The next result is a criterion for the positivity of λ_0, and is a particular case of Corollary 8.4 with $\mathbb{B} = L^1(\mu)$ in the next section. It is not deduced from the last section in terms of duality (Theorem 7.1) but conversely, it provides an improvement of Theorem 6.2 as shown by the proof of Corollary 6.6 below.

Corollary 7.3 (Criterion and basic estimates). *Without condition (7.2), we have* $\kappa^{-1}/4 \leqslant \lambda_0 \leqslant \kappa^{-1}$, *where*

$$\kappa^{-1} = \inf_{1\leqslant n\leqslant m<N+1} \biggl[\biggl(\sum_{i=1}^{n} \frac{1}{\mu_i a_i}\biggr)^{-1} + \biggl(\sum_{i=m}^{N} \frac{1}{\mu_i b_i}\biggr)^{-1}\biggr]\biggl(\sum_{j=n}^{m} \mu_j\biggr)^{-1}. \tag{7.5}$$

Furthermore, we have

$$\delta_L \wedge \delta_R \geqslant \kappa \geqslant \bigl(\mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1}\bigr)\bigl(\delta_L \wedge \delta_R\bigr),$$

where

$$\begin{aligned}
S &= \sum_{i=1}^{N} \frac{1}{\mu_i a_i} + \frac{1}{\mu_N b_N} \mathbb{1}_{\{N<\infty\}},\\
\delta_L &= \sup_{1\leqslant n<N+1} \sum_{i=1}^{n} \frac{1}{\mu_i a_i} \sum_{j=n}^{N} \mu_j, \qquad \delta_R = \sup_{1\leqslant m<N+1} \sum_{j=1}^{m} \mu_j \sum_{k=m}^{N} \frac{1}{\mu_k b_k}.
\end{aligned}$$

Note that $\delta_L = \delta^{(4.4)}$ and δ_R almost coincides with $\delta^{(3.1)}$, except for δ_R there is a shift of the state space. The second assertion of Corollary 7.3 means that $\lambda_0 > 0$ iff the process goes to either 0 or $N+1$ exponentially fast. This is intuitively clear

by (7.1). Obviously, we have $\lambda_0 = \kappa^{-1} = 0$ if $\sum_i \mu_i = \infty$ and $\sum_j (\mu_j a_j)^{-1} = \infty$ since then $\delta_L = \delta_R = \infty$. See also Corollary 8.6 below.

Proof of Corollary 6.6. For given rates (a_i, b_i) in the setup of Section 6, by (5.3), we have

$$\sum_{i=p}^{q} \hat{\nu}_i = \frac{1}{b_0}\sum_{i=p}^{q} \mu_{i-1} = \frac{1}{b_0}\sum_{i=p-1}^{q-1} \mu_i, \qquad \sum_{j=p}^{q} \hat{\mu}_j = b_0 \sum_{j=p}^{q} \nu_{j-1} = b_0 \sum_{j=p-1}^{q-1} \nu_j$$

and

$$\sum_{i=m}^{N} \frac{1}{\hat{\mu}_i \hat{b}_i} = \sum_{i=m}^{N-1} \hat{\nu}_{i+1} + \frac{1}{\hat{\mu}_N \hat{b}_N}\mathbb{1}_{\{N<\infty\}} = \frac{1}{b_0}\sum_{i=m}^{N-1} \mu_i + \frac{1}{b_0}\mu_N \mathbb{1}_{\{N<\infty\}} = \frac{1}{b_0}\sum_{i=m}^{N} \mu_i.$$

Regarding the process studied in Corollary 7.3 as a dual of the one given in the last section and then add a hat to each quantity of Corollary 7.3. It follows that

$$\begin{aligned}
\hat{\kappa}^{-1} &= \inf_{1\leqslant n\leqslant m<N+1} \bigg[\bigg(\sum_{i=1}^{n} \hat{\nu}_i\bigg)^{-1} + \bigg(\sum_{i=m}^{N} \frac{1}{\hat{\mu}_i \hat{b}_i}\bigg)^{-1}\bigg]\bigg(\sum_{j=n}^{m} \hat{\mu}_j\bigg)^{-1} \\
&= \inf_{1\leqslant n\leqslant m<N+1} \bigg[\bigg(\sum_{i=0}^{n-1} \mu_i\bigg)^{-1} + \bigg(\sum_{i=m}^{N} \mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=n-1}^{m-1} \nu_j\bigg)^{-1} \\
&= \inf_{0\leqslant n<m<N+1} \bigg[\bigg(\sum_{i=0}^{n} \mu_i\bigg)^{-1} + \bigg(\sum_{i=m}^{N} \mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=n}^{m-1} \nu_j\bigg)^{-1} \\
&= \kappa^{-1}.
\end{aligned}$$

Next, we have

$$\hat{a}_1 \widehat{S} = b_0 \bigg(\sum_{i=1}^{N} \hat{\nu}_i + \frac{1}{\hat{\mu}_N \hat{b}_N}\mathbb{1}_{\{N<\infty\}}\bigg) = \sum_{i=0}^{N} \mu_i = Z,$$

$\hat{\delta}_L = \delta_R$, and $\hat{\delta}_R = \delta_L$. Since $\sum_i \mu_i < \infty$, by Theorem 7.1 (1), we have $\lambda_1 = \hat{\lambda}_0$. Thus, Corollary 6.6 now follows from Corollary 7.3 immediately except for a slight change of the lower bound in the second assertion. For which, since $Z < \infty$, the term "$\wedge \delta_R$" is not needed (cf. Proof of Corollary 8.4). $\square$

Proof of Theorem 1.5. (a) Condition (1.3) implies that $N = \infty$ and furthermore the uniqueness of the symmetric process on $L^2(\mu)$ by Proposition 1.3. Now, $\alpha^* = \lambda_1$ or λ_0 by [2; Theorem 5.3] or Proposition 1.2, respectively.

(b) In the case that $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i b_i)^{-1} = \infty$, the process is zero-recurrent and so we have $\alpha^* = 0$. Noting that $\delta^{(3.1)}$, $\delta^{(4.4)}$, $\kappa^{(6.13)}$, and $\kappa^{(7.3)}$ are all equal to infinity, the conclusions of the theorem become obvious. Hence, in what follows, we may assume that only one of $\sum_i \mu_i$ and $\sum_i (\mu_i b_i)^{-1}$ is equal to infinity.

(c) Let $b_0 = 0$. Then the basic estimate follows from Corollary 7.3.

(d) We now prove the first two parts of the theorem under the assumption that $b_0 = 0$. In the case that $\sum_i \mu_i < \infty$ but $\sum_i (\mu_i a_i)^{-1} = \infty$, we have $\kappa^{(7.5)} = \delta^{(4.4)}$ which gives us part (1) of the theorem. Next, if $\sum_i \mu_i = \infty$ but $\sum_i (\mu_i a_i)^{-1} < \infty$, then for δ_L and δ_R given in Corollary 7.3, we have $\delta_L = \delta^{(4.4)} = \infty$ and then $\kappa^{(7.5)} < \infty$ iff $\delta_R < \infty$. Clearly, $\delta_R < \infty$ iff $\delta^{(3.1)} < \infty$ since $\sum_i (\mu_i a_i)^{-1} < \infty$. This gives us part (2) of the theorem.

(e) Finally, let $b_0 > 0$. This is a dual case of that $b_0 = 0$ treated in (c) and (d). By exchanging the measures μ and ν, we obtain the remaining conclusions of the theorem.

Actually, part (1) of the theorem is a combination of Theorems 4.2 and 6.2, and part (2) is a combination of Theorems 3.1 and 7.1. □

We are now ready to prove an extension of Theorem 1.5.

Theorem 7.4 (Criterion and basic estimates). *Without condition (1.3), Theorem 1.5 remains true provided*

(1) *the process in part (1) is replaced by the maximal one and α^* is replaced by $\alpha^{\max}$: the largest ε such that*
$$\sum_j |p_{ij}(t) - \pi_j| \leqslant C_i\, e^{-\varepsilon t}, \qquad t \geqslant 0, \tag{7.6}$$
for some $L^1(\pi)$-locally integrable function C_i depending on i only; and

(2) *the process in part (2) needs no change (i.e., the minimal one).*

Proof. Since λ_1 is equivalent to $\lambda_0^{(4.2)}$ (Theorem 6.2) and by duality, $\lambda_1 = \lambda_0^{(7.1)}$ and $\lambda_0^{(4.2)} = \lambda_0^{(2.2)}$, it is clear that $\lambda_0^{(7.1)}$ is equivalent to $\lambda_0^{(2.2)}$. Alternatively, one can use Corollary 7.3 to arrive at the same conclusion. Now, part (2) of the theorem follows by Proposition 1.2 for which we do not assume (1.3). As mentioned in the last proof, part (1) with the original α^* also follows by [2] provided (1.3) holds.

Even though in the previous study ([12], for instance), we consider only the ergodic processes under (1.2), but λ_1 can be actually identified with some exponentially ergodic convergence rate for more general ergodic processes (reversible Markov chains, in particular). First, the fact that the L^2-exponential convergence rate is described by λ_1 does not require the regularity of the Dirichlet form (cf. proof of Proposition 1.1, for instance). Next, for a Markov process with state space $(X, \mathscr{X}, \pi)$ and transition probabilities $\{P_t(x, \cdot)\}$, let $\tilde{\varepsilon}_1$ be the largest ε such that
$$\|P_t(x, \cdot) - \pi\|_{\mathrm{Var}} \leqslant C(x)\, e^{-\varepsilon t}, \qquad x \in X,\ t \geqslant 0, \tag{7.7}$$
for some $L^1(X, \pi)$-locally integrable function $C(x)$ depending on x only. Then for a reversible process having density $p_t(x, y)$, we have $\lambda_1 = \tilde{\varepsilon}_1$ provided
$$p_s(\cdot, \cdot) \in L^{1/2}_{\mathrm{loc}}(X, \pi) \qquad \text{for some } \ s > 0, \tag{7.8}$$
and the set of bounded functions with compact support is dense in $L^2(X, \pi)$. The outline of the proof is as follows.

(i) Prove that $\tilde{\varepsilon}_1 \geqslant \lambda_1$.

(ii) Show that $\|(P_t-\pi)f\|^2 \leqslant C_f\, e^{-\tilde{\varepsilon}_1 t}$ for bounded f with compact support.
(iii) Remove the constant C_f in the last line for fixed f.
(iv) Extend f to $L^2(X,\pi)$ and then claim that $\lambda_1 \geqslant \tilde{\varepsilon}_1$.

By assumption, the last step is obvious. The detailed proof for the first three steps is given, respectively, in [12]: (8.6), the last formula in §8.3 replacing ε_1 with $\tilde{\varepsilon}_1$, and the proof of Lemma 8.12. Actually, this is a small correction to [12; Theorem 8.13 (4)] (i.e., replacing ε_1 by $\tilde{\varepsilon}_1$) and its proof. It is known that $\tilde{\varepsilon}_1 > 0$ iff $\varepsilon_1 > 0$ (as well as $\varepsilon_2 > 0$ used in the original proof of the cited theorem). Hence, the exponential ergodicity is kept but the rates $\varepsilon_1 \geqslant \tilde{\varepsilon}_1 \geqslant \varepsilon_2$ may be different. By the way, we mention that the change of topology is necessary in many cases. For instance, the pointwise convergence is natural in the discrete case but is not in the continuous case. In the ergodic situation, the total variation norm is good enough in general but it is meaningless in the non-ergodic case.

Having this result at hand, part (1) of the theorem follows since we have $\alpha^{\max} = \tilde{\varepsilon}_1$ in the present context. $\square$

We now introduce an interpretation, similar to Section 5, of the duality used in Theorem 7.1. For the ergodic process with Q-matrix,

$$Q = \begin{pmatrix} -b_0 & b_0 & 0 & 0 \\ a_1 & -a_1-b_1 & b_1 & 0 \\ 0 & a_2 & -a_2-b_2 & b_2 \\ 0 & 0 & a_3 & -a_3 \end{pmatrix}, \qquad a_i, b_i > 0,$$

we have a simpler transformation matrix

$$M = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 \\ 0 & \mu_1 & \mu_2 & \mu_3 \\ 0 & 0 & \mu_2 & \mu_3 \\ 0 & 0 & 0 & \mu_3 \end{pmatrix} \Longrightarrow M^{-1} = \begin{pmatrix} \mu_0^{-1} & -\mu_0^{-1} & 0 & 0 \\ 0 & \mu_1^{-1} & -\mu_1^{-1} & 0 \\ 0 & 0 & \mu_2^{-1} & -\mu_2^{-1} \\ 0 & 0 & 0 & \mu_3^{-1} \end{pmatrix}.$$

Then

$$\begin{aligned} MQM^{-1} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ b_0 & -a_1-b_0 & a_1 & 0 \\ 0 & b_1 & -a_2-b_1 & a_2 \\ 0 & 0 & b_2 & -a_3-b_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \hat{a}_1 & -\hat{a}_1-\hat{b}_1 & \hat{b}_1 & 0 \\ 0 & \hat{a}_2 & -\hat{a}_2-\hat{b}_2 & \hat{b}_2 \\ 0 & 0 & \hat{a}_3 & -\hat{a}_3-\hat{b}_3 \end{pmatrix}. \end{aligned}$$

We obtain a process having an absorbing state at 0 and being killed at the state 3. The original trivial eigenvalue with non-zero constant eigenfunction is transferred into the trivial one with eigenfunction $\mathbb{1}_{\{0\}}$. Our dual matrix $\widehat{Q}$ is now obtained by eliminating the first row and the first column from the matrix on the right-hand side. The elimination is to make the symmetrizability of $\widehat{Q}$ and at the same time removes the trivial eigenvalue of the last matrix. Unlike the example given in Section 5 where the size of the state space stays the same: $\{0,1,2,3\} \to \{1,2,3,4\}$ with a shift for the dual one, here the size of the state space is reduced by one: $\{0,1,2,3\} \to \{1,2,3\}$.

We are now ready to examine some examples.

Examples 7.5. (1) Let $N=1$. Then the Q-matrix is degenerated to be a single killing $-c$ and so $\lambda_0=c=\kappa^{-1}$.

(2) Let $N=2$. Then

$$\lambda_0=\frac{1}{2}\Big(a_1+a_2+b_1+b_2-\sqrt{(a_1-a_2+b_1-b_2)^2+4a_2b_1}\,\Big).$$

The next two examples are taken from Chen, Zhang and Zhao (2003, Examples 2.2 and 2.3)

Examples 7.6. (1) Let $N=2$, $a_1=a_2=1$, $b_1=2$, and $b_2=3$. Then $\lambda_0=2$, and by Corollary 7.2, we have

$$\bar\delta_1\leqslant\lambda_0^{-1}=0.5\leqslant\delta_1,$$

where

$$\delta_1=\frac{4+\sqrt{3}}{10}\approx 0.573,\qquad \bar\delta_1=\frac{7}{15}=0.4\dot{6},\qquad \frac{\delta_1}{\bar\delta_1}\approx 1.23.$$

Next, $\kappa\leqslant\lambda_0^{-1}\leqslant 4\kappa$ with $\kappa=3/7$. Obviously, $\big(\bar\delta_1,\delta_1\big)\subset(\kappa,4\kappa)$.

(2) Let $N=2$, $b_1=1$, $b_2=2$,

$$a_1=\frac{2-\varepsilon^2}{1+\varepsilon},\qquad \varepsilon\in\big[0,\sqrt{2}\,\big),$$

and $a_2=1$. Then $\lambda_0=2-\varepsilon$, and we have

$$\bar\delta_1\leqslant\lambda_0^{-1}=(2-\varepsilon)^{-1}\leqslant\delta_1,$$

where

$$\begin{aligned}\delta_1&=\frac{4+\sqrt{2}+(2+\sqrt{2}\,)\varepsilon-\varepsilon^2}{8+2\,\varepsilon-3\varepsilon^2}=\frac{1}{\lambda_0}+\frac{(1+\varepsilon)(\sqrt{2}-\varepsilon)}{8+2\,\varepsilon-3\varepsilon^2},\\ \bar\delta_1&=\frac{8+6\,\varepsilon-\varepsilon^2}{16+4\,\varepsilon-6\varepsilon^2}=\frac{1}{\lambda_0}-\frac{\varepsilon^2}{2(8+2\,\varepsilon-3\varepsilon^2)}.\end{aligned}$$

Hence,

$$\frac{\delta_1}{\bar\delta_1}=2-\frac{2(4-\sqrt{2}\,)(1+\varepsilon)}{8+6\,\varepsilon-\varepsilon^2}<1.354.$$

Next, $\kappa\leqslant\lambda_0^{-1}\leqslant 4\kappa$ with

$$\kappa=\frac{1}{\lambda_0}-\min\bigg\{\frac{1}{8+2\,\varepsilon-3\,\varepsilon^2},\ \frac{\varepsilon^2}{8-4\,\varepsilon^2+\varepsilon^3}\bigg\}.$$

Even though it is not so obvious now but we do have $\big(\bar\delta_1,\delta_1\big)\subset(\kappa,4\kappa)$.

Examples 7.7. Because of Theorem 7.1, we can now transfer [10; Examples 9.27] into the present context, see Table 7.1, by using (5.1) and (5.7). Here, for the sixth example, we need a restriction: $1/k < b/a \leqslant k/(k-1)^2$ $(k \geqslant 2)$.

Table 7.1 Exact λ_0 for nine examples

$\boldsymbol{a_i}\ (i \geqslant 1)$	$\boldsymbol{b_i}\ (i \geqslant 1)$	$\boldsymbol{\lambda_0}$	$\boldsymbol{v_i}\ (i \geqslant 1)$
a	$b\ (a<b)$	$\big(\sqrt{a}-\sqrt{b}\,\big)^2$	$\sqrt{a/b}$
$\gamma_1(i-1)+\gamma_0$ $\gamma_0>0, \gamma_1\geqslant 0$	$\beta_1 i\ (\beta_1>\gamma_1)$	$\beta_1-\gamma_1$	$\dfrac{\gamma_1 i+\gamma_0}{\beta_1 i}$
$i-1+\beta_0$ $\beta_0>0$	$2(i+1)+\beta_0$	2	$\dfrac{(i+1)(i+\beta_0)}{i[2(i+1)+\beta_0]}$
i	$2i+4+\sqrt{2}$	3	$\dfrac{i+1}{2i+4+\sqrt{2}}\left[1+\dfrac{2(i+\sqrt{2}\,)}{i(i+2\sqrt{2}-1)}\right]$
$\dfrac{a}{i}$	b	$b-\dfrac{\sqrt{a^2+4ab}-a}{2}$	$\dfrac{\sqrt{a^2+4ab}+a}{2b\,i}$
a	$(i\wedge k)b$	$\big(\sqrt{bk}-\sqrt{a}\,\big)^2$	$\dfrac{1}{i\wedge k}\sqrt{ak/b}$
$i+1$	i^2	2	i^{-1}
$(i-1)^2\ (i\geqslant 2)$ $a_1>0$	i^2	$\dfrac{1}{4}$	$\dfrac{2i+1}{2(i+1)}$
$2+(-1)^{i-1}$ $(i\geqslant 2)$ $a_1=\dfrac{7-\sqrt{33}}{2}$	$2[2+(-1)^i]$	$6-\sqrt{33}$	$\dfrac{\sqrt{33}+(-1)^i}{8}$

We now go to the second part of this section. Consider the birth–death processes with a more general state space $E=\{i: -M-1<i<N+1\}$, M, $N \leqslant \infty$ and with Dirichlet boundaries at $-M-1$ if $M>-\infty$ and at $N+1$ if $N<\infty$. Its Q-matrix now is $q_{i,i+1}=b_i>0$, $q_{i,i-1}=a_i>0$, and $q_{ij}=0$ if $|i-j|>1$ for $i,j\in E$. Fix a reference point $\theta\in E$. Define

$$\begin{aligned}
\mu_{\theta+n} &= \frac{a_{\theta-1}a_{\theta-2}\cdots a_{\theta+n+1}}{b_\theta b_{\theta-1}\cdots b_{\theta+n}}, \qquad -M-1-\theta<n\leqslant -2,\\
\mu_{\theta-1} &= \frac{1}{b_\theta b_{\theta-1}}, \qquad \mu_\theta=\frac{1}{a_\theta b_\theta}, \qquad \mu_{\theta+1}=\frac{1}{a_\theta a_{\theta+1}},\\
\mu_{\theta+n} &= \frac{b_{\theta+1}b_{\theta+2}\cdots b_{\theta+n-1}}{a_\theta a_{\theta+1}\cdots a_{\theta+n}}, \qquad 2\leqslant n<N+1-\theta.
\end{aligned} \tag{7.9}$$

Correspondingly,

$$D(f)=\sum_{-M-1<i\leqslant\theta}\mu_i a_i(f_i-f_{i-1})^2+\sum_{\theta\leqslant i<N+1}\mu_i b_i(f_{i+1}-f_i)^2, \tag{7.10}$$
$$f\in\mathscr{K},\ f_{-M-1}=0 \text{ if } M<\infty \text{ and } f_{N+1}=0 \text{ if } N<\infty.$$

Let us begin with a particular application of Corollary 8.4 to $\mathbb{B}=L^1(\mu)$.

Corollary 7.8 (Criterion and basic estimates). *Let λ_0 be defined by* (7.1) *with the present state space E. Then we have $\kappa^{-1}/4 \leqslant \lambda_0 \leqslant \kappa^{-1}$, where*

$$\kappa^{-1} = \inf_{m,n\in E:\, m\leqslant n} \bigg[\bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=n}^{N} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg]\bigg(\sum_{j=m}^{n} \mu_j\bigg)^{-1}. \tag{7.11}$$

By the way, we extend Corollary 6.6 to the present general state space.

Corollary 7.9 (Criterion and basic estimates). *Let $\sum_{i\in E}\mu_i < \infty$ and define λ_1 as in (6.1). Then we have $\kappa^{-1}/4 \leqslant \lambda_1 \leqslant \kappa^{-1}$, where*

$$\kappa^{-1} = \inf_{m,n\in E:\, m< n} \bigg[\bigg(\sum_{i=-M}^{m} \mu_i\bigg)^{-1} + \bigg(\sum_{i=n}^{N} \mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=m}^{n-1} \frac{1}{\mu_j b_j}\bigg)^{-1}. \tag{7.12}$$

Proof. When $M<\infty$, the corollary is simply a modification of Corollary 6.6 by shifting the left end-point of the state space from 0 to $-M$. Thus, when $M=\infty$, we can choose a sequence $\{M_p\}_{p=1}^{\infty}$ such that $M_p \uparrow \infty$ as $p\uparrow\infty$ and then the assertion holds if M is replaced by M_p for each p. In which case, the corresponding λ_1 is denoted by $\lambda_1^{(M_p)}$ for a moment. Because $\sum_{i\in E}\mu_i<\infty$, following the proof above (4.2), it follows that

$$\lambda_1 = \inf\big\{D(f): \mu(f^2)=1,\ \mu(f)=0,\ f_i = f_{(i\vee m)\wedge n} \text{ for some } m,n\in E,\ m<n\big\}.$$

Hence, we have $\lambda_1^{(M_p)} \downarrow \lambda_1$ as $p\uparrow\infty$. Similarly, replacing M by M_p, we have the notation $\kappa^{(M_p)}$. The proof will be done once we show that

$$\big(\kappa^{(M_p)}\big)^{-1} \downarrow \kappa^{-1} \qquad \text{as } p\uparrow\infty.$$

Obviously, we have

$$\big(\kappa^{(M_p)}\big)^{-1}\downarrow \quad \text{as } p\uparrow \quad \text{and} \quad \big(\kappa^{(M_p)}\big)^{-1} \geqslant \kappa^{-1}.$$

To prove the required assertion, let $\varepsilon>0$. Then by definition of κ there exist $m_0,\ n_0\in E$, $m_0<n_0$ such that

$$\bigg[\bigg(\sum_{i=-M}^{m_0} \mu_i\bigg)^{-1} + \bigg(\sum_{i=n_0}^{N} \mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=m_0}^{n_0-1} \frac{1}{\mu_j b_j}\bigg)^{-1} \leqslant \kappa^{-1}+\varepsilon.$$

Next, since $\sum_i \mu_i<\infty$, for fixed m_0, n_0 and large enough $M_p\,(-M_p<m_0)$, we have

$$\bigg[\bigg(\sum_{i=-M_p}^{m_0} \mu_i\bigg)^{-1} + \bigg(\sum_{i=n_0}^{N} \mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=m_0}^{n_0-1} \frac{1}{\mu_j b_j}\bigg)^{-1} \leqslant \kappa^{-1}+2\,\varepsilon.$$

Combining these facts with the definition of $\kappa^{(M_p)}$, we obtain

$$\begin{aligned}\kappa^{-1} &\leqslant \big(\kappa^{(M_p)}\big)^{-1}\\ &= \inf_{-M_p\leqslant m<n<N+1}\bigg[\bigg(\sum_{i=-M_p}^{m}\mu_i\bigg)^{-1}+\bigg(\sum_{i=n}^{N}\mu_i\bigg)^{-1}\bigg]\bigg(\sum_{j=m}^{n-1}\frac{1}{\mu_j b_j}\bigg)^{-1}\\ &\leqslant \kappa^{-1}+2\,\varepsilon.\end{aligned}$$

Since ε is arbitrary, we have proved that $\big(\kappa^{(M_p)}\big)^{-1}\to\kappa^{-1}$ as $p\to\infty$. $\square$

For the remainder of this section, we study a splitting technique. It provides a different tool to study the problem having bilateral Dirichlet boundaries. This approach is especially meaningful if the duality discussed in Section 5 does not work, such as in studying the processes on the whole $\mathbb{Z}$ or the Poincaré-type inequalities given in the next section. We remark that Corollaries 7.8 and 7.9 use slightly the splitting idea only (cf. Proof (b) of Theorem 8.2 below). The idea is splitting the state space into two parts and then estimating the first (non-trivial) eigenvalue in terms of the local ones. We have used this technique several times before: Chen and Wang (1998) with Dirichlet boundary for the unbounded region, Chen, Zhang and Zhao (2003), as well as Mao and Xia (2009), with Neumann boundary. The first and the third papers work on a very general setup. Here, we follow the second one with some addition.

To state our result, we need to construct two birth–death processes on the left- and the right-hand sides, respectively, for a given birth–death process with rates (a_i, b_i) and state space E. Fix a constant $\gamma>1$.

(L) The process on the left-hand side has state space $E^{\theta-}=\{i: -M-1< i\leqslant\theta\}$, reflects at θ. Its transition structure is the same as the original one except a_θ is replaced by γa_θ.

(R) The process on the right-hand side has state space $E^{\theta+}=\{i:\theta\leqslant i< N+1\}$, reflects at θ. Its transition structure is again the same as the original one except b_θ is replaced by $\gamma(\gamma-1)^{-1}b_\theta$.

For the process on the right-hand side, the state θ is a Neumann boundary. At $N+1$, it is a Dirichlet boundary if $N<\infty$. For this process, the first eigenvalue, denoted by $\lambda_0^{\theta+,\gamma}$, has already been studied in Sections 2 and 3. With a change of the order of the state space, it follows that the process on the left-hand side has the same boundary condition, denote by $\lambda_0^{\theta-,\gamma}$ its first eigenvalue. Note that ignoring a finite number of the states does not change the positivity of $\lambda_0^{\theta\pm,\gamma}$, in the qualitative case, we simply denote them by $\lambda_0^{(\pm)}$, respectively. In general, according to $\sum_{i=\theta}^N(\mu_i b_i)^{-1}$ and/or $\sum_{i=-M}^{\theta}(\mu_i a_i)^{-1}$, $\sum_{i=\theta}^N\mu_i$ and/or $\sum_{i=-M}^{\theta}\mu_i$ being finite or not, there are eight cases for the processes on $\mathbb{Z}$. For instance, if $\sum_{i=\theta}^N(\mu_i b_i)^{-1}=\infty$, then $\lambda_0^{(+)}=0$ by Theorem 3.1. Since in this section, we are working on bilateral Dirichlet boundaries, it is natural to assume that $\lambda_0^{(\pm)}>0$. The other cases may be treated in a parallel way. For instance, when $\lambda_0^{(-)}=0$, it is more natural to consider the process on $[-M, N+1)$ with reflecting at some finite $-M$ and then pass to the limit as $M\to\infty$ (cf. the proof of Corollary 7.8).

In this case, the eigenfunction should be strictly decreasing once $\lambda_0 > 0$. Hence, there is no reason to use the splitting technique. Note that the explicit criterion for $\lambda_0^{(\pm)} > 0$ is given by Theorem 3.1. We can now state the main result of the second part of this section as follows.

Theorem 7.10.

(1) *In general, the Dirichlet eigenvalue λ_0 of the birth–death process on $E = \{i : -M-1 < i < N+1\}$ satisfies*

$$\inf_{\theta\in E}\inf_{\gamma>1}\big(\lambda_0^{\theta-,\gamma}\vee\lambda_0^{\theta+,\gamma}\big) = \lambda_0 \geqslant \sup_{-M-1\leqslant\theta\leqslant N+1}\sup_{\gamma>1}\big(\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\big), \tag{7.13}$$

where on the right-hand, when $\theta = -M-1$, define $\lambda_0^{\theta-,\gamma} = \infty$, and $\lambda_0^{\theta+,\gamma}$ to be the first eigenvalue of the original process (independent of γ) reflected at $-M$ if $M < \infty$; when $\theta = N+1$, define $\lambda_0^{\theta+,\gamma} = \infty$, and $\lambda_0^{\theta-,\gamma}$ to be the one reflected at N if $N < \infty$.

(2) *The second equality in (7.13) replacing* $\sup_{-M-1\leqslant\theta\leqslant N+1}$ *by* $\sup_{\theta\in E}$ *also holds provided $\lambda_0^{(\pm)} > 0$, and moreover,*

$$\sum_{i=-M}^{\theta}\mu_i = \infty \ \ \text{if } M = \infty \quad \text{and} \quad \sum_{i=\theta}^{N}\mu_i = \infty \ \ \text{if } N = \infty. \tag{7.14}$$

Theorem 7.10 was proved in Chen, Zhang and Zhao (2003) for the half-space (i.e., one of M and N is finite), under the hypotheses that $\sum_i(\mu_i b_i)^{-1} < \infty$ and $\sum_i \mu_i < \infty$ which is essentially the case of having a finite state space.

To prove Theorem 7.10, we need some preparation. First, we couple these two processes on a common state space $\overline{E} = \{i : -M-1 < i < N+2\}$. Next, separate the two processes by shifting the state space $E^{\theta+}$ by one to the right: $1 + E^{\theta+}$. Denote by $(\bar a_i, \bar b_i)$ the rates of the connected process. For this, we need to build a bridge for the processes on the two sides by adding two more rates $\bar b_\theta = \gamma - 1$ and $\bar a_{\theta+1} = 1$. The construction here will become clear once we have a deeper understanding about the eigenfunction and it will be explained in Part II of the proof of the theorem. Roughly speaking, there are two possible shapes of the eigenfunction, the construction enables us to transform one of them to the other so that the splitting with Neumann boundaries becomes practical. For which, one needs the parameter γ as shown in Lemma 7.12 below. In detail, we now have

$$\bar a_i = \begin{cases} a_i, & -M-1 < i \leqslant \theta-1,\\ \gamma a_\theta, & i=\theta,\\ 1, & i=\theta+1,\\ a_{i-1}, & \theta+2\leqslant i < N+2,\end{cases} \qquad \bar b_i = \begin{cases} b_i, & -M-1 < i \leqslant \theta-1,\\ \gamma-1, & i=\theta,\\ \dfrac{\gamma b_\theta}{\gamma-1}, & i=\theta+1,\\ b_{i-1}, & \theta+2\leqslant i < N+2.\end{cases}$$

Applying (7.9) to the present setup and removing the factor $b_\theta(1-\gamma)^{-1}$ (which simplifies the notation but does not change the ratio $\overline{D}(f)/\bar\mu(f^2)$), we obtain

$$\begin{aligned} &\bar\mu_i = \mu_i, && -M-1 < i \leqslant \theta-1,\\ &\bar\mu_\theta = \frac{1}{\gamma}\,\mu_\theta, && \bar\mu_{\theta+1} = \frac{\gamma-1}{\gamma}\,\mu_\theta,\\ &\bar\mu_i = \mu_{i-1}, && \theta+2\leqslant i < N+2. \end{aligned} \tag{7.15}$$

Then

$$\bar{\mu}_i \bar{a}_i = \mu_i a_i, \quad i \leqslant \theta, \quad \bar{\mu}_\theta \bar{b}_\theta = \frac{\gamma - 1}{\gamma} \mu_\theta, \quad \bar{\mu}_i \bar{b}_i = \mu_{i-1} b_{i-1}, \quad i \geqslant \theta + 1. \tag{7.16}$$

The next two results are basic in using the splitting technique.

Lemma 7.11. *Given f on E, define $\bar{f}$ on $\overline{E}$ as follows: $\bar{f}_i = f_i$ for $i \leqslant \theta$ and $\bar{f}_i = f_{i-1}$ for $i \geqslant \theta + 1$. Then we have $\bar{\mu}(\bar{f}^2) = \mu(f^2)$ and $\overline{D}(\bar{f}) = D(f)$.*

Proof. Clearly, we have $\bar{f}_\theta = \bar{f}_{\theta+1}$. Then

$$\begin{aligned}
\bar{\mu}(\bar{f}^2) &= \sum_{-M-1<i\leqslant\theta-1} \mu_i f_i^2 + (\bar{\mu}_\theta + \bar{\mu}_{\theta+1}) f_\theta^2 + \sum_{\theta+2\leqslant i<N+2} \mu_{i-1} f_{i-1}^2 = \mu(f^2), \\
\overline{D}(\bar{f}) &= \sum_{-M-1<i\leqslant\theta} \bar{\mu}_i \bar{a}_i (\bar{f}_i - \bar{f}_{i-1})^2 + \sum_{\theta+1\leqslant i<N+2} \bar{\mu}_i \bar{b}_i (\bar{f}_{i+1} - \bar{f}_i)^2 \qquad (7.17) \\
&= \sum_{-M-1<i\leqslant\theta} \mu_i a_i (f_i - f_{i-1})^2 + \sum_{\theta+1\leqslant i<N+2} \mu_{i-1} b_{i-1} (f_i - f_{i-1})^2 \\
&= D(f) \qquad (\text{by } (7.10)). \quad \square
\end{aligned}$$

Lemma 7.12. *For a given birth–death process with state space E and rates (a_i, b_i), if its eigenfunction g of λ satisfies $g_{\theta-1} < g_\theta > g_{\theta+1}$ (resp. $g_{\theta-1} > g_\theta < g_{\theta+1}$) for some $\theta \in E$ (of course, $g_{-M-1} = 0$ if $M < \infty$, and $g_{N+1} = 0$ if $N = \infty$), let*

$$\gamma = 1 + \frac{b_\theta (g_\theta - g_{\theta+1})}{a_\theta (g_\theta - g_{\theta-1})} > 1, \tag{7.18}$$

and let $\bar{g}_i = g_i$ for $i \leqslant \theta$, $\bar{g}_i = g_{i-1}$ for $i \geqslant \theta + 1$. Then for the $(\bar{a}_i, \bar{b}_i)$-process, $\bar{g}$ is the eigenfunction of $\bar{\lambda} = \lambda$ having the property $\bar{g}_{\theta+1} = \bar{g}_\theta$. Furthermore, $\bar{g}|_{(-M-1,\theta]}$ is the eigenfunction of $\bar{\lambda}$ of the process on the left-hand side reflecting at θ, and similarly $\bar{g}|_{[\theta+1,N+1)}$ is the eigenfunction of the process on the right-hand side reflecting at $\theta + 1$.

Proof. By the construction of $(\bar{a}_i, \bar{b}_i)$ and $\bar{g}$, we have

$$\overline{\Omega}\, \bar{g}(i) = \begin{cases} \Omega g(i) = -\lambda g_i = -\bar{\lambda} \bar{g}_i, & i \leqslant \theta - 1, \\ \Omega g(i-1) = -\lambda g_{i-1} = -\bar{\lambda} \bar{g}_i, & i \geqslant \theta + 2. \end{cases}$$

Next, by (7.15), we have

$$\begin{aligned}
\overline{\Omega}\, \bar{g}(\theta) = \bar{b}_\theta (\bar{g}_{\theta+1} - \bar{g}_\theta) + \bar{a}_\theta (\bar{g}_{\theta-1} - \bar{g}_\theta) = \bar{a}_\theta (g_{\theta-1} - g_\theta) = \gamma a_\theta (g_{\theta-1} - g_\theta), \\
\overline{\Omega}\, \bar{g}(\theta+1) &= \bar{b}_{\theta+1} (\bar{g}_{\theta+2} - \bar{g}_{\theta+1}) + \bar{a}_{\theta+1} (\bar{g}_\theta - \bar{g}_{\theta+1}) \\
&= \bar{b}_{\theta+1} (g_{\theta+1} - g_\theta) \\
&= \frac{\gamma}{\gamma - 1} b_\theta (g_{\theta+1} - g_\theta).
\end{aligned}$$

In the first formula, the term containing $\bar{b}_\theta$ vanishes. This is the reason why we can regard θ as a reflecting boundary for the process on the left-hand side.

Similarly, one can regard $\theta+1$ as the one for the process on the right-hand side in view of the second formula. By (7.18), the right-hand sides are the same which is equal to

$$\bigg[1+\frac{b_\theta(g_\theta-g_{\theta+1})}{a_\theta(g_\theta-g_{\theta-1})}\bigg]a_\theta(g_{\theta-1}-g_\theta)=a_\theta(g_{\theta-1}-g_\theta)+b_\theta(g_\theta-g_{\theta+1})=-\lambda g_\theta=-\bar\lambda\bar g_\theta.$$

We have thus proved the lemma. $\square$

Proof of Theorem 7.10. *Part I.* In this part, we prove Theorem 7.10 (1) with the first "=" replaced by "$\geqslant$". The proof of this part is relatively easier. Let $f\in\mathscr{K}$, $f\neq 0$. Define $\bar f$ as in Lemma 7.11. For fixed $\theta\in E$ and $\gamma>1$, noting that re-labeling the state space does not change $\lambda_0^{\theta+,\gamma}$, by (7.17), we have

$$\overline{D}(\bar f)\geqslant\lambda_0^{\theta-,\gamma}\sum_{i\leqslant\theta}\bar\mu_i\bar f_i^2+\lambda_0^{\theta+,\gamma}\sum_{i\geqslant\theta+1}\bar\mu_i\bar f_i^2\geqslant\big(\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\big)\,\bar\mu\big(\bar f^2\big).$$

Hence by Lemma 7.11,

$$\frac{D(f)}{\mu(f^2)}=\frac{\overline{D}(\bar f)}{\bar\mu\big(\bar f^2\big)}\geqslant\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}.$$

Making the supremum with respect to γ and θ, it follows that

$$\frac{D(f)}{\mu(f^2)}\geqslant\sup_{-M-1<\theta<N+1}\ \sup_{\gamma>1}\big(\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\big).$$

At the boundaries, say $\theta=-M-1$ for instance, by (7.10) and the convention, we have

$$D(f)\geqslant\sum_{-M-1<i<N+1}\mu_i b_i(f_{i+1}-f_i)^2\geqslant\lambda_0^{\theta+,\gamma}\,\mu\big(f^2\big)=\big[\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\big]\mu\big(f^2\big).$$

Therefore, we indeed have

$$\frac{D(f)}{\mu(f^2)}\geqslant\sup_{-M-1\leqslant\theta\leqslant N+1}\ \sup_{\gamma>1}\big(\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\big).$$

Making infimum with respect to f, we obtain

$$\lambda_0=\inf_{f\in\mathscr{K},\,f\neq 0}\frac{D(f)}{\mu(f^2)}\geqslant\sup_{-M-1\leqslant\theta\leqslant N+1}\ \sup_{\gamma>1}\big(\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\big).$$

This proves the (second) inequality in (7.13).

To prove the upper estimate, fix $\theta\in E$ and $\gamma>1$ again. As we have seen from the last part of proof (g) of Theorem 2.4 and Proposition 2.5, if we let $\lambda_0^{\theta+,\gamma,n}$ denote the local eigenvalue with Neumann boundary at θ and Dirichlet boundary at $n+1$, then $\lambda_0^{\theta+,\gamma,n}\downarrow\lambda_0^{\theta+,\gamma}$ as $n\uparrow\infty$. Thus, for each $\varepsilon>0$, we

have $\lambda_0^{\theta+,\gamma,n} < \lambda_0^{\theta+,\gamma} + \varepsilon$ for large enough n. By Proposition 2.2, we can assume that the corresponding eigenfunction $g^{(+,n)}$ of $\lambda_0^{\theta+,\gamma,n}$ satisfies $g_\theta^{(+,n)} = 1$ and $g_i^{(+,n)} = 0$ for all $i > n\,(> \theta)$. Similarly, we have $\lambda_0^{\theta-,\gamma,m} < \lambda_0^{\theta-,\gamma} + \varepsilon$ for small enough $-m$, and moreover, the eigenfunction $g^{(-,m)}$ of $\lambda_0^{\theta-,\gamma,m}$ satisfies $g_\theta^{(-,m)} = 1$ and $g_i^{(-,m)} = 0$ for all $i < -m\,(< \theta)$. Let $\bar f$ be defined as above, connecting $g^{(-,m)}$ and $g^{(+,n)}$. Then $\bar f$ has a finite support, $\bar f_\theta = \bar f_{\theta+1} = 1$, and moreover by (7.10),

$$\begin{aligned}\overline{D}(\bar f) &= \sum_{\overline{E}\ni i\leqslant\theta} \bar\mu_i\bar a_i\big(\bar f_i - \bar f_{i-1}\big)^2 + \sum_{\overline{E}\ni i\geqslant\theta+1} \bar\mu_i\bar b_i\big(\bar f_{i+1} - \bar f_i\big)^2\\ &= \lambda_0^{\theta-,\gamma,m}\sum_{i\leqslant\theta}\bar\mu_i\bar f_i^2 + \lambda_0^{\theta+,\gamma,n}\sum_{i\geqslant\theta+1}\bar\mu_i\bar f_i^2\\ &\leqslant \big(\lambda_0^{\theta-,\gamma}\vee\lambda_0^{\theta+,\gamma}+\varepsilon\big)\,\bar\mu\big(\bar f^2\big).\end{aligned}$$

By Lemmas 7.11 and 7.12, this gives us

$$\lambda_0 = \bar\lambda_0 \leqslant \lambda_0^{\theta-,\gamma}\vee\lambda_0^{\theta+,\gamma}$$

since ε is arbitrary. Furthermore, we have

$$\lambda_0 \leqslant \inf_{\theta\in E}\,\inf_{\gamma>1}\big(\lambda_0^{\theta-,\gamma}\vee\lambda_0^{\theta+,\gamma}\big)$$

as required. □

The proof of the equalities in Theorem 7.10 is much harder. For which, we need once again a deeper understanding of the eigenfunction of λ_0. To have a concrete impression, we mention that the eigenfunction in Examples 7.6 (2) is $g_0 = g_3 = 0$, $g_1 = (1+\varepsilon)g_2$. Thus, when $\varepsilon = 0$, we have $g_1 = g_2$. Besides, it is rather easy to see the shape of eigenfunction g of the examples given in Table 7.1 since $v_i < 1$ iff $g_{i+1} < g_i$ for all i.

Definition 7.13.

(1) *A function f is said to be unimodal if there exists a finite k such that f_i is strictly increasing for $i \leqslant k$ and strictly decreasing for $i \geqslant k$.*
(2) *A function f is said to be a simple echelon if there exists a k such that $f_k = f_{k+1}$, f_i is strictly increasing for $i \leqslant k$ and strictly decreasing for $i \geqslant k+1$.*

Proposition 7.14. *Let g be a positive eigenfunction of $\lambda > 0$ for a birth–death process. Then g is strictly monotone, or unimodal, or a simple echelon.*

Proof. (a) Let $g_k \geqslant g_{k+1}$ for some k. We prove that g is strictly decreasing for $i \geqslant k+1$. To do so, note that

$$b_{k+1}(g_{k+2} - g_{k+1}) = -\lambda g_{k+1} - a_{k+1}(g_k - g_{k+1}) \leqslant -\lambda g_{k+1} < 0.$$

Thus, we have $g_{k+2} < g_{k+1}$. Assume that $g_n < g_{n-1}$ for some $n \geqslant k+2$. Then the eigenequation shows that

$$b_n(g_{n+1} - g_n) = -\lambda g_n - a_n(g_{n-1} - g_n) < -\lambda g_n < 0.$$

By induction, this gives us $g_{n+1} < g_n$ for all $n \geqslant k+1$.

(b) By symmetry, we can handle with the case that $g_k \leqslant g_{k+1}$ for some k. One starts at

$$a_k(g_{k-1} - g_k) = -\lambda g_k - b_k(g_{k+1} - g_k) \leqslant -\lambda g_k < 0.$$

We obtain $g_{k-1} < g_k$ and then $g_{n-1} < g_n$ for all $n \leqslant k$ by induction.

(c) By (a) and (b), it follows that there is no local convex part of g. Otherwise, there is a k such that either $g_{k-1} > g_k < g_{k+1}$ or $g_{k-1} > g_k = g_{k+1} < g_{k+2}$ which contradict what we proved in (a) and (b).

(d) We claim that for every k, say $k = 0$ for simplicity, the two cases "$g_{-1} \geqslant g_0$" and "$g_0 \leqslant g_1$" cannot happen at the same time. Otherwise, there are four situations:

$$g_{-1} = g_0 = g_1, \quad g_{-1} > g_0 < g_1, \quad g_{-1} > g_0 = g_1, \quad \text{and} \quad g_{-1} = g_0 < g_1.$$

The first one cannot happen, otherwise we have $g_i \equiv 0$. By (c), the second case is impossible. The last two cases are also impossible by (b) and (a), respectively.

(e) Having these preparations at hand, we are ready to prove the main assertion of the proposition. Clearly, we need only to consider the case that g is not strictly monotone. Choose a starting point, say 0 for instance. By (d), we have only one possibility: either $g_{-1} \geqslant g_0$ or $g_0 \leqslant g_1$. Without loss of generality, assume that $g_0 \leqslant g_1$. If $g_0 = g_1$, then by (a) and (b), g is a simple echelon. If $g_0 < g_1$, then on the one hand, by (b), g_i is strictly increasing for all $i \leqslant 1$, and on the other hand, we can find a $k \geqslant 1$ such that $g_1 < g_2 < \ldots < g_k \geqslant g_{k+1}$ since g is not strictly monotone by assumption. Applying (a) again, it follows that g is strictly decreasing for all $i \geqslant k+1$. Hence, g is either unimodal or a simple echelon. $\square$

Proposition 7.15. *For the birth–death process on $\mathbb{Z}$, the following assertions hold.*

(1) *The eigenfunction g of λ satisfies the following successive formulas:*

$$\begin{aligned} g_{k+1} &= g_k + \frac{1}{\mu_k b_k}\left[\mu_\theta a_\theta(g_\theta - g_{\theta-1}) - \lambda \sum_{i=\theta}^{k} \mu_i g_i\right], \qquad k \geqslant \theta, \\ g_{k-1} &= g_k + \frac{1}{\mu_k a_k}\left[\mu_\theta a_\theta(g_{\theta-1} - g_\theta) - \lambda \sum_{i=k}^{\theta-1} \mu_i g_i\right], \qquad k < \theta. \end{aligned} \tag{7.19}$$

(2) *If $\lambda = 0$, then the non-trivial eigenfunction g with $g_\theta = 1$ for some $\theta \in \mathbb{Z}$ is given by*

$$\begin{aligned} g_n &= 1 + (1 - g_{\theta-1}) \sum_{j=\theta}^{n-1} \prod_{k=\theta}^{j} \frac{a_k}{b_k}, \qquad n \geqslant \theta, \\ g_n &= 1 - (1 - g_{\theta-1}) \sum_{j=n}^{\theta-1} \prod_{k=j+1}^{\theta-1} \frac{b_k}{a_k}, \qquad n < \theta. \end{aligned}$$

In this case, the function g is either the constant function $\mathbb{1}$ or strictly monotone on $\mathbb{Z}$.

(3) *If $\lambda > 0$ and*

$$\sum_{i=-\infty}^{\theta} \mu_i = \sum_{i=\theta}^{\infty} \mu_i = \infty, \tag{7.20}$$

then the non-trivial eigenfunction g of λ cannot be monotone.

Proof. (a) Part (1) of the proposition follows from the eigenequation.

(b) When $\lambda = 0$, with $u_i := g_{i+1} - g_i$ $(i \in \mathbb{Z})$, the eigenequation $b_i u_i = a_i u_{i-1}$ gives us

$$u_j = (1 - g_{\theta-1}) \prod_{k=\theta}^{j} \frac{a_k}{b_k}, \quad j \geqslant \theta, \qquad u_j = (1 - g_{\theta-1}) \prod_{k=j+1}^{\theta-1} \frac{b_k}{a_k}, \quad j < \theta.$$

It follows that either $g_i \equiv 1$ or g is strictly monotone on $\mathbb{Z}$. Now, part (2) of the proposition follows by making a summation of j from θ to $n-1$ and from n to $\theta - 1$, respectively.

(c) Without loss of generality, assume that $g_\theta = 1$ for some $\theta \in \mathbb{Z}$. Suppose that g is non-decreasing, then by the first equation in part (1), we would have

$$\infty > \frac{\mu_\theta a_\theta (g_\theta - g_{\theta-1})}{\lambda} \geqslant \sum_{k=\theta}^{n} \mu_k g_k \geqslant \sum_{k=\theta}^{n} \mu_k \to \infty \qquad \text{as } n \to \infty.$$

Otherwise, if g is non-increasing, then by the second equation in part (1), we would have

$$\infty > \frac{\mu_\theta a_\theta (g_{\theta-1} - g_\theta)}{\lambda} \geqslant \sum_{k=n}^{\theta-1} \mu_k g_k \geqslant \sum_{k=n}^{\theta-1} \mu_k \to \infty \qquad \text{as } n \to -\infty.$$

We have thus proved part (3) of the proposition. $\square$

We remark that Proposition 7.15 (2) is different from Proposition 2.2 where the eigenfunction of $\lambda = 0$ must be a constant. Here is a simple example with $\theta = 0$: $a_i = b_i = |i|$ if $i \neq 0$ and $a_0 = b_0 = 1$, then corresponding to $\lambda = 0$, we have a family of linear eigenfunctions $\{g_i^{(\gamma)} = 1 + (1-\gamma)\, i : i \in \mathbb{Z}\}_{\gamma \in \mathbb{R}}$ (normalized at 0) with one-parameter γ.

Proposition 7.16. *Let* (7.14) *hold and g be a non-zero eigenfunction of $\lambda_0 > 0$. Then g is either positive or negative on E.*

Proof. If one of M or N is finite, then the conclusion follows from Proposition 2.2 (1). From now on in the proof, assume that $M = N = \infty$.

(a) If the conclusion of the proposition does not hold, then there is a k (say) such that $g_k \leqslant 0$ and either $g_{k-1} > 0$ or $g_{k+1} > 0$. By symmetry, assume that $g_{k+1} > 0$.

(b) We now prove that $g_i > 0$ for all $i \geqslant k+1$. Given $m, n \in \mathbb{Z}$ with $m \leqslant n$, denote by $\lambda_0^{[m,n]}$ the first eigenvalue of the process restricted on the state space $\{i : m \leqslant i \leqslant n\}$ with Dirichlet boundaries at $m-1$ and $n+1$ in the sense similar to (7.1). If the assertion does not hold, then there is a $k_0 : k_0 > k+1$ such that $g_{k_0} \leqslant 0$. Now, let $\tilde{g}$ satisfy $\tilde{g}_k = 0$, $\tilde{g}_i = g_i$ for $i = k+1, \ldots, k_0 - 1$, $\tilde{g}_{k_0} = \varepsilon$ for some $\varepsilon > 0$, $\tilde{g}_i = 0$ for $i \geqslant k_0 + 1$. Note that

$$\begin{aligned}\big(-\Omega\,\tilde{g}\big)(k+1) &= b_{k+1}\big(\tilde{g}_{k+1} - \tilde{g}_{k+2}\big) + a_{k+1}\big(\tilde{g}_{k+1} - \tilde{g}_k\big)\\ &= b_{k+1}(g_{k+1} - g_{k+2}) + a_{k+1}(g_{k+1} - g_k) + a_{k+1} g_k\\ &= \lambda_0 g_{k+1} + a_{k+1} g_k\\ &\leqslant \lambda_0\, \tilde{g}_{k+1},\end{aligned}$$

Because of $\lambda_0 > 0$ and following proof (b) of Proposition 2.1, we can choose a suitable $\varepsilon > 0$ such that

$$\sum_{i=k+1}^{k_0} \mu_i \tilde{g}_i\big(-\Omega\tilde{g}\big)(i) < \lambda_0 \sum_{i=k+1}^{k_0} \mu_i \tilde{g}_i^2.$$

It follows that $\lambda_0^{[k+1,k_0]} < \lambda_0$. However, it is obvious that $\lambda_0 \leqslant \lambda_0^{[k+1,k_0]}$ and so we get a contradiction. We have thus proved that $g_i > 0$ for all $i \geqslant k+1$.

(c) By (7.14) and proof (c) of Proposition 7.15, g cannot be non-decreasing since $\lambda_0 > 0$. Hence, there is a $\theta \geqslant k+2$ such that $g_{k+1} < g_{k+2} < \ldots < g_\theta \geqslant g_{\theta+1}$. In the case that $g_\theta > g_{\theta+1}$, by introducing an additional point but keeping the same λ_0 as shown in Lemma 7.12, one can reduce to the case that $g_\theta = g_{\theta+1}$. Hence, one can split the original process into two as in (L) and (R). Now, starting from θ at which $g_\theta > 0$, look at the process on the left-hand side in the inverse way, one finds the point $k < \theta$ at which $g_k \leqslant 0$. Applying proof (b) above to this process, one may get a contradiction. It follows that $g > 0$ on $(-\infty, \theta] \supset (-\infty, k]$.

Therefore, we should have $g > 0$ on $\mathbb{Z}$. $\square$

Proof of Theorem 7.10. *Part II.* We now prove the equalities in (7.13). By assumption $\lambda_0^{(\pm)} > 0$ and the second inequality in (7.13), it follows that $\lambda_0 > 0$. If one of M and N is finite, then the non-trivial eigenfunction g must be positive by Proposition 2.2 (1). In this case, it is helpful to include the boundary into the domain of g for understanding its shape. Then by Proposition 7.14, there are only two possibilities:

(i) g is unimodal;
(ii) g is a simple echelon.

Next, if $M = N = \infty$, then by Proposition 7.16, we have $g > 0$. Moreover, by Proposition 7.15, g cannot be monotone. Hence, by Proposition 7.14, g has again one of shapes (i) and (ii) as above.

We now prove the equalities in (7.13) only in the case that $M = N = \infty$. The proof for the other case is simpler.

(a) Case (ii). We use the operator $I\!I$ defined in Section 2:

$$\overline{I\!I}_i^{\theta+,\gamma}(\bar{f}) = \frac{1}{\bar{f}_i} \sum_{j=i}^{N+1} \frac{1}{\bar{\mu}_j \bar{b}_j} \sum_{k=\theta+1}^{j} \bar{\mu}_k \bar{f}_k, \qquad \theta + 1 \leqslant i < N+2.$$

For each $\bar{f}$ satisfying: $\bar{f}_i = f_i$ for $i \leqslant \theta$ and $\bar{f}_i = f_{i-1}$ for $i \geqslant \theta+1$ for some f on E, by (7.15) and (7.16), we have

$$\begin{aligned}
\overline{II}_i^{\theta+,\gamma}(\bar{f}) &= \frac{1}{f_{i-1}} \sum_{j=i}^{N+1} \frac{1}{\mu_{j-1} b_{j-1}} \left[\bar{\mu}_{\theta+1} f_\theta + \sum_{k=\theta+2}^{j} \mu_{k-1} f_{k-1} \right] \\
&= \frac{1}{f_{i-1}} \sum_{j=i}^{N+1} \frac{1}{\mu_{j-1} b_{j-1}} \left[\left(1 - \frac{1}{\gamma}\right) \mu_\theta f_\theta + \sum_{k=\theta+1}^{j-1} \mu_k f_k \right] \\
&= \frac{1}{f_{i-1}} \sum_{j=i-1}^{N} \frac{1}{\mu_j b_j} \sum_{k=\theta}^{j} \mu_k f_k - \frac{\mu_\theta f_\theta}{\gamma f_{i-1}} \sum_{j=i-1}^{N} \frac{1}{\mu_j b_j} \\
&= \frac{1}{f_{i-1}} \sum_{j=i-1}^{N} \frac{1}{\mu_j b_j} \sum_{k=\theta+1}^{j} \mu_k f_k + \left(1 - \frac{1}{\gamma}\right) \frac{\mu_\theta f_\theta}{f_{i-1}} \sum_{j=i-1}^{N} \frac{1}{\mu_j b_j} \\
&\qquad \theta+1 \leqslant i < N+2. \qquad (7.21)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\overline{II}_i^{\theta-,\gamma}(\bar{f}) &= \frac{1}{\bar{f}_i} \sum_{j=-M}^{i} \frac{1}{\bar{\mu}_j \bar{a}_j} \sum_{k=j}^{\theta} \bar{\mu}_k \bar{f}_k \\
&= \frac{1}{f_i} \sum_{j=-M}^{i} \frac{1}{\mu_j a_j} \sum_{k=j}^{\theta} \mu_k f_k - \left(1 - \frac{1}{\gamma}\right) \frac{\mu_\theta f_\theta}{f_i} \sum_{j=-M}^{i} \frac{1}{\mu_j a_j}, \\
&\qquad -M-1 < i \leqslant \theta. \qquad (7.22)
\end{aligned}$$

Because $g_\theta = g_{\theta+1}$, we can regard θ as a Neumann boundary of the original process restricted on the left–hand side and at the same time, regard θ as a Neumann boundary of the original process restricted on the right–hand side. Because $\lambda_0^{(\pm)} > 0$, by Proposition 2.5 (2), we have $g_{\pm\infty} = 0$. Hence, by (2.11), (7.21), and (7.22), we obtain

$$\overline{II}_i^{\theta+,\gamma}(\bar{g}) = \frac{1}{\lambda_0} + \left(1 - \frac{1}{\gamma}\right) \frac{\mu_\theta g_\theta}{g_{i-1}} \sum_{j=i-1}^{N} \frac{1}{\mu_j b_j}, \qquad \theta+1 \leqslant i < N+2,$$

$$\overline{II}_i^{\theta-,\gamma}(\bar{g}) = \frac{1}{\lambda_0} - \left(1 - \frac{1}{\gamma}\right) \frac{\mu_\theta g_\theta}{g_i} \sum_{j=-M}^{i} \frac{1}{\mu_j a_j}, \qquad -M-1 < i \leqslant \theta.$$

By Proposition 2.2 (2), we have

$$\sup_{\theta \leqslant i < N+1} \frac{\mu_\theta g_\theta}{g_i} \sum_{j=i}^{N} \frac{1}{\mu_j b_j} \leqslant \frac{1}{\lambda_0}, \qquad \sup_{-M-1 < i \leqslant \theta} \frac{\mu_\theta g_\theta}{g_i} \sum_{j=-M}^{i} \frac{1}{\mu_j a_j} \leqslant \frac{1}{\lambda_0}.$$

Therefore, by the second inequality in (7.13) and Theorem 2.4 (3), it follows that

$$\begin{aligned}
\lambda_0 &\geqslant \sup_{\theta'\in E}\ \sup_{\gamma>1}\Big[\lambda_0^{\theta'-,\gamma}\wedge\lambda_0^{\theta'+,\gamma}\Big]\\
&\geqslant \sup_{\gamma>1}\Big[\lambda_0^{\theta-,\gamma}\wedge\lambda_0^{\theta+,\gamma}\Big]\\
&\geqslant \sup_{\gamma>1}\Big[\Big(\inf_{-M-1<i\leqslant\theta}\overline{I\!I}_i^{\theta-,\gamma}(\bar g)^{-1}\Big)\wedge\Big(\inf_{\theta+1\leqslant i<N+2}\overline{I\!I}_i^{\theta+,\gamma}(\bar g)^{-1}\Big)\Big]\\
&= \sup_{\gamma>1}\ \inf_{\theta+1\leqslant i<N+2}\overline{I\!I}_i^{\theta+,\gamma}(\bar g)^{-1}\\
&= \sup_{\gamma>1}\bigg\{\frac{1}{\lambda_0}+\Big(1-\frac{1}{\gamma}\Big)\sup_{\theta\leqslant i<N+1}\frac{\mu_\theta g_\theta}{g_i}\sum_{j=i}^{N}\frac{1}{\mu_j b_j}\bigg\}^{-1}\\
&= \lambda_0.
\end{aligned}$$

We have thus proved in Case (ii) the second equality in (7.13).

To prove the first equality in (7.13), noting the inequality was proved in Part I, we have dually

$$\begin{aligned}
\lambda_0 &\leqslant \inf_{\theta'\in E}\ \inf_{\gamma>1}\Big[\lambda_0^{\theta'-,\gamma}\vee\lambda_0^{\theta'+,\gamma}\Big]\\
&\leqslant \inf_{\gamma>1}\Big[\lambda_0^{\theta-,\gamma}\vee\lambda_0^{\theta+,\gamma}\Big]\\
&\leqslant \inf_{\gamma>1}\Big[\Big(\sup_{-M-1<i\leqslant\theta}\overline{I\!I}_i^{\theta-,\gamma}(\bar g)^{-1}\Big)\vee\Big(\sup_{\theta+1\leqslant i<N+2}\overline{I\!I}_i^{\theta+,\gamma}(\bar g)^{-1}\Big)\Big]\\
&= \inf_{\gamma>1}\ \sup_{-M-1<i\leqslant\theta}\overline{I\!I}_i^{\theta-,\gamma}(\bar g)^{-1}\\
&= \sup_{\gamma>1}\bigg\{\frac{1}{\lambda_0}-\Big(1-\frac{1}{\gamma}\Big)\sup_{-M-1<i\leqslant\theta}\frac{\mu_\theta g_\theta}{g_i}\sum_{j=-M}^{i}\frac{1}{\mu_j a_j}\bigg\}^{-1}\\
&= \lambda_0.
\end{aligned}$$

However, there is a problem in the second line of the proof. To apply Theorem 2.4 (3), one requires that either $g\in L^2(\mu)$ or g is local. Hence, an additional work is required. Anyhow, the conclusion holds whenever both M and N are finite. We will come back to the proof in proof (c) below.

(b) Case (i). By Lemma 7.12, this case can be reduced to Case (ii). Actually, the proof becomes easier now. With γ given by (7.18), we have

$$\begin{aligned}
\overline{I\!I}_i^{\theta+,\gamma}(\bar g)&\equiv\frac{1}{\lambda_0}, \qquad \theta+1\leqslant i<N+2,\\
\overline{I\!I}_i^{\theta-,\gamma}(\bar g)&\equiv\frac{1}{\lambda_0}, \qquad -M-1<i\leqslant\theta.
\end{aligned}$$

Hence the second equality in (7.13) holds. Moreover, the first equality in (7.13) also holds whenever both M and N are finite.

(c) To complete the proof for the first equality in (7.13), we need to overcome the unbounded problem. For this, choose $M_p, N_p \uparrow \infty$ as $p \to \infty$. Denote by $\lambda_0^{\theta-,\gamma,p}$, $\lambda_0^{\theta+,\gamma,p}$ and $\lambda_0^{(p)}$, respectively, the quantities $\lambda_0^{\theta-,\gamma}$, $\lambda_0^{\theta+,\gamma}$, and λ_0 when M and N are replaced by M_p and N_p. Note that for a finite state space, we certainly have $\lambda_0^{(p)} > 0$, its eigenfunction is positive (by Proposition 2.2 (1)) and has properties (i) and (ii) mentioned in the above proof (by Proposition 7.14). Clearly, for each fixed θ and γ, we have

$$\lambda_0^{\theta\pm,\gamma,p} \downarrow \lambda_0^{\theta\pm,\gamma}, \quad \lambda_0^{(p)} \downarrow \lambda_0 \quad \text{as } p \to \infty.$$

Thus, as proved in (a) and (b), whether we are in Case (i) or (ii), we have for each p,

$$\begin{aligned}\lambda_0^{(p)} &= \inf_{\theta\in[-M_p,N_p]} \inf_{\gamma>1} \big[\lambda_0^{\theta-,\gamma,p} \vee \lambda_0^{\theta+,\gamma,p}\big]\\ &\geqslant \inf_{\theta\in[-M_p,N_p]} \inf_{\gamma>1} \big[\lambda_0^{\theta-,\gamma} \vee \lambda_0^{\theta+,\gamma}\big]\\ &\geqslant \inf_{\theta\in E} \inf_{\gamma>1} \big[\lambda_0^{\theta-,\gamma} \vee \lambda_0^{\theta+,\gamma}\big].\end{aligned}$$

Therefore, by the first inequality in (7.13) proved in Part I, it follows that

$$\lambda_0 \leqslant \inf_{\theta\in E} \inf_{\gamma>1} \big[\lambda_0^{\theta-,\gamma} \vee \lambda_0^{\theta+,\gamma}\big] \leqslant \lambda_0^{(p)} \downarrow \lambda_0 \qquad \text{as } p \to \infty.$$

We have thus completed the proof of the theorem. $\square$

Here are remarks about the assumption made in part (2) of Theorem 7.10. Similar to the upper estimate, we do have

$$\lambda_0^{(p)} = \sup_{\theta\in[-M_p,N_p]} \sup_{\gamma>1} \Big[\lambda_0^{\theta-,\gamma,p} \wedge \lambda_0^{\theta+,\gamma,p}\Big].$$

The problem is that $\lambda_0^{\theta\pm,\gamma,p} \downarrow \lambda_0^{\theta\pm,\gamma}$ as $p \to \infty$ goes to the opposite direction and the approximating sequences $\{M_p\}$ and $\{N_p\}$ depend on θ and γ. Hence, the proof for the upper estimate does not work for the lower one. Next, to prove the second equality in (7.13), it seems more natural to assume that $\lambda_0^{\theta-,\gamma} \wedge \lambda_0^{\theta+,\gamma} > 0$ for some θ and γ, that is, $\lambda_0^{(-)} \wedge \lambda_0^{(+)} > 0$, rather than $\lambda_0^{(-)} \vee \lambda_0^{(+)} > 0$ as we made. However, if one of them is zero, say $\lambda_0^{(-)} = 0$, then as mentioned before Theorem 7.10, we have a single Dirichlet boundary but not the bilateral Dirichlet ones, and the variational formula takes a different form (i.e., the second inequality in (7.13) at the boundaries). Condition (7.14) is due to the same reason. In particular, when $M = -1$, for instance, if $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i a_i)^{-1} = \infty$, then $\lambda_0^{(+)} = 0$ by Theorem 3.1, and we go back to the case studied in Section 4. In which case, the eigenfunction of λ_0 is strictly increasing.

To conclude this section, we introduce a complement result to [12; Proposition 5.13] about the principal eigenvalue for general Markov chains.

Proposition 7.17. *Let $(q_{ij} : i, j \in E)$ be symmetric with respect to (μ_i) on a countable set E, not necessarily conservative (or having killings):*

$$d_i := q_i - \sum_{j \neq i} q_{ij} \geqslant 0, \qquad q_i := -q_{ii} \in [0, \infty).$$

Define

$$D(f) = \frac{1}{2} \sum_{i,j \in E} \mu_i q_{ij} (f_j - f_i)^2 + \sum_{i \in E} \mu_i d_i f_i^2$$

and

$$\lambda_0 = \inf \big\{ D(f) : f \text{ has a finite support and } \mu\big(f^2\big) = 1 \big\}.$$

Then we have $\inf_{i \in E} q_i \geqslant \lambda_0$.

Proof. Without loss of generality, assume that $E = \mathbb{Z}_+ = \{0, 1, \dots\}$. Fix $k \in E$ and take $f = \mathbb{1}_{\{k\}}$. Then $\mu\big(f^2\big) = \mu_k$ and

$$\begin{aligned} D(f) &= \sum_{i,j:\, i<j} \mu_i q_{ij} (f_j - f_i)^2 + \sum_{i \in E} \mu_i d_i f_i^2 \\ &= \sum_{j>k} \mu_k q_{kj} (f_k - f_j)^2 + \sum_{i<k} \mu_i q_{ik} (f_i - f_k)^2 + \sum_{i \in E} \mu_i d_i f_i^2 \\ &= \sum_{j>k} \mu_k q_{kj} + \sum_{i<k} \mu_i q_{ik} + \mu_k d_k. \end{aligned}$$

By the symmetry of $\mu_i q_{ij}$, we get

$$D(f) = \sum_{j>k} \mu_k q_{kj} + \sum_{i<k} \mu_k q_{ki} + \mu_k d_k = \mu_k \Bigg(\sum_{j \neq k} q_{kj} + d_k \Bigg) = \mu_k q_k.$$

It follows that

$$\lambda_0 \leqslant D(f) / \mu\big(f^2\big) = q_k.$$

The assertion now follows since $k \in E$ is arbitrary. $\square$

8. Criteria for Poincaré-type inequalities

As in [9] for the ergodic case having $N < \infty$ or (1.2), the results studied in Sections 2, 3 and 7 can be extended to a more general setup, so called Poincaré-type inequalities. In this way, one obtains various types of stability, not only the L^2-exponential one studied in the other sections of the paper. Here we consider only the criteria and the basic estimates for the inequalities. In other words, we extend Theorems 3.1, 4.2, and 6.2 to the general setup with some improvement. At the same time, we introduce a criterion for the processes studied in Section 7 in this setup. To do so, we need a class of normed linear spaces $(\mathbb{B}, \|\cdot\|_{\mathbb{B}}, \mu)$ consisting of real Borel measurable functions on a measurable space $(X, \mathscr{X}, \mu)$. We now modify the hypotheses on the normed linear spaces given in [12; Chapter 7] as follows.

(H1) In the case that $\mu(X) = \infty$, $\mathbb{1}_K \in \mathbb{B}$ for all compact K. Otherwise, $1 \in \mathbb{B}$.
(H2) If $h \in \mathbb{B}$ and $|f| \leqslant h$, then $f \in \mathbb{B}$.
(H3) $\|f\|_{\mathbb{B}} = \sup_{g \in \mathscr{G}} \int_X |f| g \mathrm{d}\mu$,

where $\mathscr{G}$, to be specified case by case, is a class of nonnegative $\mathscr{X}$-measurable functions. A typical example is $\mathscr{G} = \{\mathbb{1}\}$ and then $\mathbb{B} = L^1(\mu)$. Throughout this section, we assume (H1)–(H3) for $(\mathbb{B}, \|\cdot\|_{\mathbb{B}}, \mu)$ without mentioning again.

Before moving further, let us mention the following result.

Remark 8.1. *Without using (1.2), the results in [9] remain true under the condition $\sum_i \mu_i < \infty$ replacing the original process with the maximal one if necessary.*

The key reason is that without condition (1.2), the same conclusion holds in Section 4 on which the cited paper is based on.

In this section, our state space is $E = \{i : -M-1 < i < N+1\}$ $(M, N \leqslant \infty)$ as in the second part of Section 7. The next result is the main one in this section; it has several corollaries as we have seen in the last section. Note that the factor 4 in (8.2) below is universal, independent of $\mathbb{B}$.

Theorem 8.2. *Consider the minimal birth–death process with Dirichlet boundaries at $-M-1$ if $M < \infty$ and at $N+1$ if $N < \infty$. Assume that $\mathscr{G}$ contains a locally positive element. Then the optimal constant $A_{\mathbb{B}}$ in the Poincaré-type inequality*

$$\big\|f^2\big\|_{\mathbb{B}} \leqslant A_{\mathbb{B}} D(f), \qquad f \in \mathscr{D}^{\min}(D), \tag{8.1}$$

satisfies

$$B_{\mathbb{B}} \leqslant A_{\mathbb{B}} \leqslant 4B_{\mathbb{B}}, \tag{8.2}$$

where the isoperimetric constant $B_{\mathbb{B}}$ can be expressed as follows:

$$B_{\mathbb{B}}^{-1} = \inf_{m,n\in E:\, m\leqslant n} \bigg[\bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=n}^{N} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg] \|\mathbb{1}_{[m,n]}\|_{\mathbb{B}}^{-1}. \tag{8.3}$$

In particular, when $\mathbb{B} = L^{p/2}(\mu)\,(p \geqslant 2)$ (then (8.1) is called the Sobolev-type inequality), we have

$$B_p^{-1} = \inf_{m,n\in E:\, m\leqslant n} \bigg[\bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=n}^{N} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg]\bigg(\sum_{j=m}^{n} \mu_j\bigg)^{-2/p}. \tag{8.4}$$

Proof. (a) First consider the transient case, in particular when one of M or N is finite. We use the proof of [11; Corollary 4.1] or [12; Corollary 7.5] with a slight modification. In proof (a) there, it was shown that one can replace "$f|_K \geqslant 1$" by "$f|_K = 1$" in computing the capacity $\mathrm{Cap}(K)$ for compact K. Without loss of generality, assume that $f \geqslant 0$. Otherwise, replace f with $|f|$. In the proof just mentioned, the condition "$\sum_i \mu_i < \infty$" was used so that $\mathbb{1} \in \mathscr{D}(D)$. We cannot use this assumption now, but for a given nonnegative $f \in \mathscr{D}^{\min}(D) \cap \mathscr{C}_c(E)$, where $\mathscr{C}_c(E)$ is the set of continuous functions with compact support, we can simply choose a nonnegative smooth $h \in \mathscr{C}_c(E)$ such that $h|_{\mathrm{supp}\,(f)} = 1$. Then $h \in \mathscr{D}^{\min}(D)$, $f \wedge h \in \mathscr{D}^{\min}(D)$, and so one can use $f \wedge h \in \mathscr{D}^{\min}(D)$ instead of $f \wedge \mathbb{1}$ to arrive at the same conclusion $D(f) \geqslant D(f \wedge h)$ as in the original proof (a).

The first step in the original proof (b) shows that one can replace a finite number of disjointed finite intervals $\{K_i\}$ by the connected one $[\min \cup_i K_i,\ \max \cup_i K_i]$. This part of the proof needs no change.

Note that in the original proof, the state space is $\{1, 2, \ldots\}$ with Dirichlet boundary at 0. The main body of the original proof (b) is to find a minimizer (actually unique) $f \in \mathscr{C}_c(E)$ for $D(f)$ having the properties $f_0 = 0$ and $f|_K = 1$. Replacing N with q for the consistence with the notation used here and let $K = \{m, m+1, \ldots, n\}$ ($1 \leqslant m \leqslant n$, here m and n are exchanged from the original proof). Now, within the class of f: $f_0 = 0$, $f|_K = 1$ and $\operatorname{supp}(f) = \{1, \ldots, q\}$ ($n \leqslant q < N+1$), the minimal solution is

$$D(f) = \bigg(\sum_{i=1}^{m} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=n}^{q} \frac{1}{\mu_i b_i}\bigg)^{-1}. \tag{8.5}$$

To handle with the general state space, one needs to move the original left-end point 1 of the state space to somewhere, say $p > -M-1$. In detail, replace the condition $m \geqslant 1$ used in defining the compact set K by $m > -M-1$. At the same time, replace $\{1, \ldots, q\}$ by $\{p, p+1, \ldots, q\}$ with $-M-1 < p \leqslant m$ for the $\operatorname{supp}(f)$. Then the last formula reads as follows:

$$D(f) = \bigg(\sum_{i=p}^{m} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=n}^{q} \frac{1}{\mu_i b_i}\bigg)^{-1}.$$

In the original proof, the ergodic condition and (1.2) are mainly used here to remove the second term on the right-hand side. We now keep it. Since the right-hand side is increasing in p and decreasing in q, by making the infimum with respect to f, it follows that

$$\begin{aligned}\operatorname{Cap}(K) &:= \inf\big\{D(f) : f \in \mathscr{D}^{\min} \cap \mathscr{C}_c(E) \text{ and } f|_K \geqslant 1\big\}\\ &= \bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=n}^{N} \frac{1}{\mu_i b_i}\bigg)^{-1}, \quad K = \{m, m+1, \ldots, n\} =: [m, n].\end{aligned}$$

The assertion of the theorem now follows by using

$$B_{\mathbb{B}} := \sup_K \frac{\|\mathbb{1}_K\|_{\mathbb{B}}}{\operatorname{Cap}(K)} = \sup_{-M-1 < m \leqslant n < N+1} \frac{\|\mathbb{1}_{[m,n]}\|_{\mathbb{B}}}{\operatorname{Cap}([m,n])}$$

and applying [11; Theorem 1.1] or [12; Theorem 7.2]. The last result is an extension of Fukushima and Uemura (2003, Theorem 3.1).

(b) Next, consider the recurrent case: both $\sum_{i<\theta} (\mu_i a_i)^{-1}$ and $\sum_{i>\theta} (\mu_i b_i)^{-1}$ are diverged. Here is actually a direct proof of the lower estimate in (8.2). Without loss of generality, assume that the reference point $\theta = 0$. Fix $m' \geqslant m \geqslant 0$ and $n' \geqslant n \geqslant 0$. Based on the knowledge about the eigenfunction given in the last section, and similar to proof (b) of Theorem 3.1, define

$$f_i = \begin{cases} \displaystyle\sum_{k=i\vee n}^{n'} \frac{1}{\mu_k b_k} \mathbb{1}_{\{i \leqslant n'\}}, & i \geqslant 0,\\[2ex] \displaystyle\gamma \sum_{k=-m'}^{i\wedge(-m)} \frac{1}{\mu_k a_k} \mathbb{1}_{\{i \geqslant -m'\}}, & i \leqslant 0,\end{cases}$$

where

$$\gamma := \gamma(m', m, n, n') = \sum_{k=n}^{n'} \frac{1}{\mu_k b_k} \Big/ \sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k}.$$

Here, γ is chosen to make f be a constant on $[-m, n]$. By (7.10), we have

$$\begin{aligned}
D(f) &= \sum_{i=-m'}^{-m} \mu_i a_i (f_i - f_{i-1})^2 + \sum_{i=n}^{n'} \mu_i b_i (f_{i+1} - f_i)^2 \\
&= \gamma^2 \sum_{i=-m'}^{-m} \frac{1}{\mu_i a_i} + \sum_{i=n}^{n'} \frac{1}{\mu_i b_i} \\
&= \bigg(\sum_{i=n}^{n'} \frac{1}{\mu_i b_i}\bigg)\bigg[1 + \bigg(\sum_{k=n}^{n'} \frac{1}{\mu_k b_k}\bigg)\bigg(\sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k}\bigg)^{-1}\bigg].
\end{aligned}$$

Moreover,

$$\|f^2\|_{\mathbb{B}} \geqslant \|f|_{[-m,n]}^2\|_{\mathbb{B}} = \bigg(\sum_{i=n}^{n'} \frac{1}{\mu_i b_i}\bigg)^2 \|\mathbb{1}_{[-m,n]}\|_{\mathbb{B}}.$$

Hence,

$$A_{\mathbb{B}} \geqslant \frac{\|f^2\|_{\mathbb{B}}}{D(f)} \geqslant \|\mathbb{1}_{[-m,n]}\|_{\mathbb{B}} \bigg[\bigg(\sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k}\bigg)^{-1} + \bigg(\sum_{j=n}^{n'} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg]^{-1}.$$

From this, we obtain the lower estimate in (8.2). Since $\mathscr{G}$ contains a locally positive element, we have $\|\mathbb{1}_{[-m,n]}\|_{\mathbb{B}} > 0$ for large enough m and n. Letting $m', n' \to \infty$, by the recurrent assumption, it follows that $A_{\mathbb{B}} = \infty$. Besides, it is obvious that $B_{\mathbb{B}} = \infty$ in this case and so the first and then the second assertion of the theorem becomes trivial in the recurrent case. $\square$

Proof (b) above indicates an easy improvement of the lower bound of $A_{\mathbb{B}}$. Use the same f as above, and define

$$\begin{aligned}
h_i^{(m,m',n,n')} &= \bigg[1 - \sum_{k=i+1}^{-m} \frac{1}{\mu_k a_k} \Big/ \sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k}\bigg]^2 \mathbb{1}_{[-m',-m-1]}(i) \\
&\quad + \bigg[1 - \sum_{k=n}^{i-1} \frac{1}{\mu_k b_k} \Big/ \sum_{k=n}^{n'} \frac{1}{\mu_k b_k}\bigg]^2 \mathbb{1}_{[n+1,n']}(i).
\end{aligned}$$

Then a simple computation shows that

$$f^2 = \bigg(\sum_{i=n}^{n'} \frac{1}{\mu_i b_i}\bigg)^2 \big(\mathbb{1}_{[-m,n]} + h^{(m,m',n,n')}\big).$$

Hence

$$\frac{\|f^2\|_{\mathbb{B}}}{D(f)} \geqslant \big\|\mathbb{1}_{[-m,n]} + h^{(m,m',n,n')}\big\|_{\mathbb{B}} \bigg[\bigg(\sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k}\bigg)^{-1} + \bigg(\sum_{j=n}^{n'} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg]^{-1}.$$

Noting that the right-hand side is increasing in m' and n', and making a change of the variable $-m \to m$, we obtain

$$A_{\mathbb{B}} \geqslant \sup_{m,n\in E:\, m\leqslant n} \big\|\mathbb{1}_{[m,n]} + h^{(-m,M,n,N)}\big\|_{\mathbb{B}} \bigg[\bigg(\sum_{k=-M}^{m} \frac{1}{\mu_k a_k}\bigg)^{-1} + \bigg(\sum_{j=n}^{N} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg]^{-1}.$$

Denote by $C_{\mathbb{B}}$ the right-hand side. Then the conclusion of Theorem 8.2 can be restated as $B_{\mathbb{B}} \leqslant C_{\mathbb{B}} \leqslant A_{\mathbb{B}} \leqslant 4B_{\mathbb{B}}$. Certainly, this remark is meaningful in other cases but we will not mention again.

The next result is an easier consequence of Theorem 8.2.

Corollary 8.3. *Everything in the premise is the same as in Theorem 8.2. Then*

(1) *we have $B_{\mathbb{B}} \leqslant B_L \wedge B_R$, where*

$$B_L = \sup_{n\in E} \sum_{i=-M}^{n} \frac{1}{\mu_i a_i} \|\mathbb{1}_{[n,N+1)}\|_{\mathbb{B}}, \qquad B_R = \sup_{n\in E} \sum_{i=n}^{N} \frac{1}{\mu_i b_i} \|\mathbb{1}_{(-M-1,n]}\|_{\mathbb{B}}.$$

The equality sign holds once

$$S := \sum_{i=-M}^{N} \frac{1}{\mu_i a_i} + \frac{1}{\mu_N b_N} \mathbb{1}_{\{N<\infty\}} = \infty.$$

(2) *Next, we have $B_{\mathbb{B}} \geqslant (B_L \wedge B_R)\mathbb{1}_{\{S=\infty\}} + S^{-1} B$, where*

$$B = \sup_{m,n\in E:\, m\leqslant n} \bigg[\bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i}\bigg)\bigg(\sum_{k=n}^{N} \frac{1}{\mu_k b_k}\bigg) \|\mathbb{1}_{[m,n]}\|_{\mathbb{B}}\bigg].$$

Proof. Clearly, by (8.3), we have

$$B_{\mathbb{B}}^{-1} \geqslant \inf_{m\leqslant n} \bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i} \|\mathbb{1}_{[m,n]}\|_{\mathbb{B}}\bigg)^{-1} = \inf_{m\in E} \bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i} \|\mathbb{1}_{[m,N+1)}\|_{\mathbb{B}}\bigg)^{-1},$$

and so $B_{\mathbb{B}} \leqslant B_L$. The equality sign holds once $\sum_{i=\theta}^{N} (\mu_i b_i)^{-1} = \infty$. Similarly, we have $B_{\mathbb{B}} \leqslant B_R$. The equality sign holds once $\sum_{i=-M}^{\theta} (\mu_i a_i)^{-1} = \infty$. Hence, $B_{\mathbb{B}} \leqslant B_L \wedge B_R$ and the equality sign holds once $S = \infty$.

Next, when $S < \infty$, we have

$$B_{\mathbb{B}}^{-1} \leqslant S \inf_{m\leqslant n} \bigg[\bigg(\sum_{i=-M}^{m} \frac{1}{\mu_i a_i}\bigg)\bigg(\sum_{k=n}^{N} \frac{1}{\mu_k b_k}\bigg) \|\mathbb{1}_{[m,n]}\|_{\mathbb{B}}\bigg]^{-1} = S B^{-1}.$$

We have thus proved the corollary. $\square$

Of course, one can decompose the constant B in Corollary 8.3 (2). For instance, for fixed m_0, we have

$$B \geqslant \bigg(\sum_{i=-M}^{m_0} \frac{1}{\mu_i a_i}\bigg) \sup_{m_0 \leqslant n < N+1} \bigg[\bigg(\sum_{k=n}^{N} \frac{1}{\mu_k b_k}\bigg) \|\mathbb{1}_{[m_0,n]}\|_{\mathbb{B}}\bigg].$$

The last factor is close to B_R when m_0 is negative enough. However, when $m_0 \to -M$, the first term tends to zero since $S < \infty$, unless $M < \infty$. This indicates that bounding $B_{\mathbb{B}}$ by B_L and B_R is rather rough, especially in the case that $E = \mathbb{Z}$ (cf. Example 8.9 below). This is a particularly different point of the processes on the whole $\mathbb{Z}$ or on the half space $\mathbb{Z}_+$ as shown by Corollary 8.4 below.

We now specify Theorem 8.2 and Corollary 8.3 to the half space: either M or N is finite. This corresponds to the processes studied in the first part of Section 7 (see Corollary 7.3).

Corollary 8.4. *In Theorem 8.2, let $M = -1$. Then we have*

$$B_{\mathbb{B}}^{-1} = \inf_{1 \leqslant n \leqslant m < N+1} \bigg[\bigg(\sum_{i=1}^{n} \frac{1}{\mu_i a_i}\bigg)^{-1} + \bigg(\sum_{i=m}^{N} \frac{1}{\mu_i b_i}\bigg)^{-1}\bigg] \|\mathbb{1}_{[n,m]}\|_{\mathbb{B}}^{-1}. \tag{8.6}$$

Furthermore, we have

$$B_L \wedge B_R \geqslant B_{\mathbb{B}} \geqslant \big(\mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1}\big)\, (B_L \wedge B_R),$$

where

$$B_L = \sup_{1 \leqslant n < N+1} \sum_{i=1}^{n} \frac{1}{\mu_i a_i} \|\mathbb{1}_{[n,N+1)}\|_{\mathbb{B}}, \qquad B_R = \sup_{1 \leqslant m < N+1} \sum_{k=m}^{N} \frac{1}{\mu_k b_k} \|\mathbb{1}_{[1,m]}\|_{\mathbb{B}},$$

$$S = \sum_{i=1}^{N} \frac{1}{\mu_i a_i} + \frac{1}{\mu_N b_N} \mathbb{1}_{\{N<\infty\}}. \tag{8.7}$$

Proof. The first assertion follows from Theorem 8.2 with $M = -1$ and an exchange of m and n again. The second one follows from Corollary 8.3 except the last estimate. When $S = \infty$, we have $B_{\mathbb{B}} = B_L$. While when $S < \infty$, we have

$$\begin{aligned} B_{\mathbb{B}}^{-1} &\leqslant S \inf_{1 \leqslant n \leqslant m < N+1} \bigg[\bigg(\sum_{i=1}^{n} \frac{1}{\mu_i a_i}\bigg)\bigg(\sum_{k=m}^{N} \frac{1}{\mu_k b_k}\bigg) \|\mathbb{1}_{[n,m]}\|_{\mathbb{B}}\bigg]^{-1} \\ &\leqslant a_1 S \inf_{1 \leqslant m < N+1} \bigg(\sum_{k=m}^{N} \frac{1}{\mu_k b_k} \|\mathbb{1}_{[1,m]}\|_{\mathbb{B}}\bigg)^{-1} \\ &= a_1 S\, B_R^{-1}. \end{aligned}$$

Therefore,

$$B_{\mathbb{B}} \geqslant B_L \mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1} B_R \geqslant \big(\mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1}\big)\, (B_L \wedge B_R)$$

as required. $\square$

When one of M or N is finite and its Dirichlet boundary is replaced by the Neumann one, the solution becomes simpler. The next result corresponds to the processes studied in Sections 2 and 3.

Theorem 8.5. *Let $M = 0$ be the Neumann boundary and assume that $\mathscr{G}$ contains a locally positive element. Then the isoperimetric constant $B_{\mathbb{B}} \coloneqq \sup_K \|\mathbb{1}_K\|_{\mathbb{B}}/\mathrm{Cap}(K)$ can be expressed as*

$$B_{\mathbb{B}} = \sup_{0\leqslant n<N+1} \sum_{i=n}^{N} \frac{1}{\mu_i b_i} \|\mathbb{1}_{[0,n]}\|_{\mathbb{B}}. \tag{8.8}$$

In particular, for the Sobolev-type inequality, we have

$$B_p = \sup_{0\leqslant n<N+1} \sum_{i=n}^{N} \frac{1}{\mu_i b_i} \bigg(\sum_{j=0}^{n} \mu_j\bigg)^{2/p}, \qquad p \geqslant 2. \tag{8.9}$$

Proof. The proof is nearly the same as that of Theorem 8.2 except one point. In proof (b) of [11; Corollary 4.1] or [12; Corollary 7.5], to find a minimizer f for $D(f)$, since the constraint $f_0 = 0$ and $f_n = 1$, f cannot be a constant on $\{0, 1, \ldots, n\}$. Now, without the constraint $f_0 = 0$, the minimizer should satisfy $f_j = 1$ for all $j : 0 \leqslant j \leqslant n$. Thus, instead of (8.5), the minimal solution becomes

$$D(f) = \bigg(\sum_{i=n}^{q} \frac{1}{\mu_i b_i}\bigg)^{-1}.$$

Then the necessary change of the proof of Theorem 8.2 after (8.5) should be clear. $\square$

Applying Theorem 8.5 to $\mathbb{B} = L^1(\mu)$, we return to Theorem 3.1. Actually, in parallel to [9], one may extend the results in Sections 2 and 3, Theorem 3.1 in particular, to the present setup of normed linear spaces and then deduce Theorem 8.5. The next result is obvious, it says that for a null-recurrent process, the L^p $(p \geqslant 1)$-Sobolev inequality is still not weak enough.

Corollary 8.6. *Consider a birth–death process on $\mathbb{Z}_+$. If $\sum_{i\geqslant 1} \mu_i = \infty$ and $\sum_{i\geqslant 1} (\mu_i b_i)^{-1} = \infty$, then $B_p^{(8.4)} = B_p^{(8.9)} = \infty$ for all $p \geqslant 2$.*

Remark 8.7. We now compare (8.6) and (8.8) in the particular case that $\sum_i \mu_i = \infty$. Then the constant $B_{\mathbb{B}}$ given in (8.6) becomes

$$B_{\mathbb{B}}^{(8.6)} = \sup_{1\leqslant m<N+1} \sum_{i=m}^{N} \frac{1}{\mu_i b_i} \|\mathbb{1}_{[1,m]}\|_{\mathbb{B}}.$$

Rewrite the constant $B_{\mathbb{B}}$ given in (8.8) as

$$B_{\mathbb{B}}^{(8.8)} = \bigg(\sum_{i=0}^{N} \frac{1}{\mu_i b_i} \|\mathbb{1}_{\{0\}]}\|_{\mathbb{B}}\bigg) \bigvee \bigg(\sup_{1\leqslant n<N+1} \sum_{i=n}^{N} \frac{1}{\mu_i b_i} \|\mathbb{1}_{[0,n]}\|_{\mathbb{B}}\bigg).$$

By (H3), we have

$$\|\mathbb{1}_{[1,n]}\|_{\mathbb{B}} \leqslant \|\mathbb{1}_{[0,n]}\|_{\mathbb{B}} \leqslant \|\mathbb{1}_{\{0\}]}\|_{\mathbb{B}} + \|\mathbb{1}_{[1,n]}\|_{\mathbb{B}}.$$

Next, by (H1), we have $\|\mathbb{1}_{\{0\}}\|_{\mathbb{B}} < \infty$. It follows that $B_{\mathbb{B}}^{(8.6)} < \infty$ iff $B_{\mathbb{B}}^{(8.8)} < \infty$.

We conclude this section by a simple example to show the role of the Poincaré-type inequalities.

Example 8.8. *Consider a birth–death process on $\mathbb{Z}_+$ with $\mu_i = (i+1)^\gamma$ $(\gamma > 1)$ and $b_i \equiv 1$. Then $a_i = i^\gamma (i+1)^{-\gamma}$ and*

$$B_p^{(8.9)} = \sup_{n \geqslant 0} \bigg[\sum_{i=0}^{n} (i+1)^\gamma\bigg]^{2/p} \sum_{j=n}^{\infty} \frac{1}{(j+1)^\gamma}, \qquad p \geqslant 2.$$

Hence, $B_p^{(8.9)} < \infty$ iff

$$p \geqslant 2\bigg(1 + \frac{2}{\gamma - 1}\bigg).$$

However, $\delta^{(3.1)} = B_2^{(8.9)} = \infty$ for all $\gamma > 1$.

Example 8.9. *Let $E = \mathbb{Z}$, $b_i \equiv 1$, $\mu_i = e^{i^2}$, and $\mathbb{B} = L^1(\mu)$. Then for the quantities given in Corollary 8.3, we have $B_L = B_R = \infty$ but $B_{\mathbb{B}} < \infty$.*

Proof. Obviously, $B_L = B_R = \infty$. To show that $B_{\mathbb{B}} < \infty$, since

$$x \vee y \leqslant x + y \leqslant 2(x \vee y),$$

it suffices to prove that

$$\sup_{m \leqslant n} \bigg[\bigg(\sum_{i=-\infty}^{m} \frac{1}{\mu_i a_i}\bigg) \bigwedge \bigg(\sum_{k=n}^{\infty} \frac{1}{\mu_k b_k}\bigg)\bigg] \sum_{j=m}^{n} \mu_j < \infty.$$

By symmetry, without loss of generality, it is enough to show that

$$\sup_{m \geqslant n \geqslant 0} \bigg(\sum_{i=-\infty}^{-m} \frac{1}{\mu_i a_i}\bigg) \sum_{j=-m}^{n} \mu_j < \infty,$$

or

$$\sup_{m \geqslant 0} \bigg(\sum_{i=-\infty}^{-m} \frac{1}{\mu_i a_i}\bigg) \sum_{j=-m}^{m} \mu_j < \infty.$$

Equivalently,

$$\sup_{m \geqslant 0} \bigg(\sum_{i=m}^{\infty} \frac{1}{\mu_i b_i}\bigg) \sum_{j=0}^{m} \mu_j < \infty. \tag{8.10}$$

The assertion now follows by using Conte's inequality:

$$x\bigg(1 + \frac{x}{24} + \frac{x^2}{12}\bigg) e^{-3x^2/4} < e^{-x^2} \int_0^x e^{y^2} \leqslant \frac{\pi^2}{8x}\big(1 - e^{-x^2}\big), \qquad x \geqslant 0$$

and Gautschi's estimate:

$$\frac{1}{2}\Big[(x^p + 2)^{1/p} - x\Big] < e^{x^p} \int_x^\infty e^{-y^p} dy \leqslant C_p\bigg[\bigg(x^p + \frac{1}{C_p}\bigg)^{1/p} - x\bigg], \qquad x \geqslant 0,$$

$$C_p = \Gamma\big(1 + 1/p\big)^{p/(p-1)}, \quad p > 1; \qquad C_2 = \pi/4.$$

Alternatively, one may check directly that the function under supremum on the left-hand side of (8.10) is decreasing in $m\,(\geqslant 1)$ and then (8.10) follows easily. $\square$

9. General killing

In Sections 4 and 7, we have studied the special case having a killing at 1 only. We now study the process with general killing, as described by (2.1) with state space shifted by 1: $E = \{i : 1 \leqslant i < N+1\}$. We use the same symmetric measure (μ_i) as in Section 4.

The next preliminary result is quite useful. To which it is more convenient to use $a_1 + c_1$ and $b_N + c_N$ for the killing rates at boundaries 1 and N (if $N < \infty$), respectively, rather than c_1 and c_N used in Proposition 2.1. Note that the killing rates in the next proposition are allowed to be zero identically.

Proposition 9.1. *Let (a_i) and (b_i) be positive but $a_1 \geqslant 0$, $b_N \geqslant 0$ if $N < \infty$, and let (c_i) be nonnegative on E. Define $\lambda_0 = \lambda_0(a_i, b_i, c_i)$ as follows:*

$$\lambda_0 = \inf\big\{D(f) : \mu\big(f^2\big) = 1, f \in \mathscr{K}\big\},$$

where

$$D(f) = \sum_{i \in E} \mu_i b_i (f_{i+1} - f_i)^2 + \mu_1 a_1 f_1^2 + \sum_{i \in E} \mu_i c_i f_i^2, \qquad f_{N+1} = 0 \text{ if } N < \infty.$$

Write $\tilde{\lambda}_0 = \lambda_0(a_i, b_i, 0)$ for simplicity. Then we have

(1) *$\lambda_0(a_i, b_i, c_i') \geqslant \lambda_0(a_i, b_i, c_i)$ if $c_i' \geqslant c_i$ for all $i \in E$.*
(2) *$\lambda_0(a_i, b_i, c_i + c) = \lambda_0(a_i, b_i, c_i) + c$ for constant $c \geqslant 0$.*
(3) *$\tilde{\lambda}_0 + \sup_{i \in E} c_i \geqslant \lambda_0 \geqslant \tilde{\lambda}_0 + \inf_{i \in E} c_i$ and the equalities hold if c_i is a constant on E.*

Proof. Since a change of $\{c_i\}_{i=1}^N$ makes no influence to $\{\mu_i\}_{i=1}^N$, part (1) is simply a comparison of the Dirichlet forms on the same space $L^2(\mu)$ with common core $\mathscr{K}$. Similarly, one can prove the other assertions. $\square$

Note that Proposition 9.1 makes a comparison for the killing rates only. Actually, a more general comparison is available in view of [3; Theorem 3.1]. Next, if (1.3) holds, then by Proposition 1.3 and the remark below (4.3), the Dirichlet is unique, and so the condition $f \in \mathscr{K}$ can be ignored in defining $\tilde{\lambda}_0$.

It is worthy to mention that the principal eigenvalue λ_0 studied here can be extended to a more general class of Schrödinger operators. That is, we may replace the nonnegative potential (c_i) with the one bounded below by a constant: $\inf_i c_i \geqslant -M > -\infty$. Then we have $c_i + M \geqslant 0$ for all i and

$$\lambda_0(a_i, b_i, c_i) = \lambda_0(a_i, b_i, c_i + M) - M \geqslant -M.$$

Having Proposition 9.1 at hand, all the examples for $\tilde{\lambda}_0$ given in Sections 3, 5 and 7, can be translated into the case of λ_0 with constant killing rate. For instance, we have the following example which already shows the complexity of the problem studied in this section.

Example 9.2. *Let $a_i \equiv a > 0$ for $i \geqslant 2$, $b_i \equiv b > 0$ and $c_i \equiv c \geqslant 0$ for $i \geqslant 1$.*

(1) *If $a_1 = a$, or $a_1 = 0$ but still $a \leqslant b$, then $\lambda_0 = \big(\sqrt{a} - \sqrt{b}\big)^2 + c$.*
(2) *If $a_1 = 0$ and $a > b$, then $\lambda_0 = c$.*

Proof. By Proposition 9.1, we need only to study $\tilde{\lambda}_0$. In the last case, since the process is ergodic, we have $\tilde{\lambda}_0 = 0$. Next, we have $\tilde{\lambda}_0 = \big(\sqrt{a} - \sqrt{b}\big)^2$ according different cases by

(i) Example 5.3 if $a_1 = a$ and $a \geqslant b$,
(ii) Example 7.7 if $a_1 = a$ and $a \leqslant b$,
(iii) Example 3.4 if $a_1 = 0$ and $a \leqslant b$. □

From now on, we return to a convention made in Section 2, the rates a_1 and b_N are combined into c_1 and c_N if $N < \infty$. Thus, in Theorem 7.1 for instance, we have $a_1 = 0$ and $c_1 > 0$, and moreover, $b_N = 0$ and $c_N > 0$ if $N < \infty$. In general, we assume that $c_i \not\equiv 0$. Otherwise, we will return to what we treated in Sections 2 and 3. Define the operator R:

$$R_i(v) = a_i\big(1 - v_{i-1}^{-1}\big) + b_i(1 - v_i) + c_i, \qquad i \in E,\ v_0 = \infty,\ v_N = 0 \text{ if } N < \infty,$$

for v in the set $\mathscr{V} = \{v_i > 0 : 1 \leqslant i < N\}$. Next, define $\widetilde{\mathscr{V}} = \mathscr{V}$ if $N < \infty$. When $N = \infty$, define

$$\begin{aligned}\widetilde{\mathscr{V}} = \bigcup_{m=1}^{\infty} &\Big\{v_i : v_i > 0 \text{ for } i < m,\ v_i = 0 \text{ for } i \geqslant m\Big\}\\ &\bigcup \Big\{v : v_i > 0 \text{ on } E, \text{ the function } f\colon f_1 = 1, f_i = \prod_{k=1}^{i-1} v_k\ (i \geqslant 2) \text{ is in } L^2(\mu)\\ &\qquad\qquad \text{and satisfies } \Omega f/f \leqslant \eta \text{ on } E \text{ for some constant } \eta\Big\}.\end{aligned} \tag{9.1}$$

For $v \in \widetilde{\mathscr{V}}$ with finite support, $R_\bullet(v)$ is also well defined by setting $1/0 = \infty$.

Theorem 9.3. *Assume that $c_i \not\equiv 0$. For λ_0 defined by (2.2) with state space $E = \{i : 1 \leqslant i < N + 1\}$, the following variational formulas hold:*

$$\inf_{v \in \widetilde{\mathscr{V}}} \sup_{i \in E} R_i(v) = \lambda_0 = \sup_{v \in \mathscr{V}} \inf_{i \in E} R_i(v). \tag{9.2}$$

Proof. (a) First, we study the lower estimate. In the case that $\sum_{k \in E} \mu_k < \infty$, as a particular consequence of [5; Theorem 1.1], we have

$$\lambda_0 \geqslant \sup_{g > 0} \inf_{i \in E} \frac{-\Omega g}{g}(i) = \sup_{g > 0} \inf_{i \in E} \bigg[a_i\bigg(1 - \frac{g_{i-1}}{g_i}\bigg) + b_i\bigg(1 - \frac{g_{i+1}}{g_i}\bigg) + c_i\bigg], \tag{9.3}$$

where $g_0 := 0$ and $g_{N+1} = 0$ if $N < \infty$. The proof remains true when $\sum_{k \in E} \mu_k = \infty$, simply using $E_m = \{1, 2, \ldots, m\}\,(m < N + 1)$ instead of the original one.

Actually, the conclusion holds in a very general setup (cf. Shiozawa and Takeda (2005) and its extension to the unbounded test functions by Zhang (2007)).

Suppose that $\lambda_0 > 0$ for a moment. Then by Proposition 2.1 (with a shift by 1 of the state space), the eigenfunction g of λ_0 is positive. It follows that the first equality sign in (9.3) can be attained and so does the last equality in (9.2) with $v_i = g_{i+1}/g_i > 0\,(1 \leqslant i < N)$. Next, if $\lambda_0 = 0$, then $N = \infty$ since a_i and b_i are positive for $i : 2 \leqslant i < N$, and $c_i \not\equiv 0$ (in the case of Theorem 7.1, we have $c_1 > 0$ and also $c_N > 0$ if $N < \infty$). By setting $v_i \equiv 1$ for $i \in E$, we get

$$\inf_{i\in E}\left[a_i\left(1-\frac{1}{v_{i-1}}\right)+b_i(1-v_i)+c_i\right] \geqslant \inf_{i\in E} c_i \geqslant 0. \tag{9.4}$$

Hence, the last term of (9.2) is nonnegative. Therefore, the last equality in (9.2) is trivial if $\lambda_0 = 0$, in view of (9.3).

(b) Next, we study the upper estimate. We consider only the case that $N = \infty$. Otherwise, the proof is easier. Given $v \in \widetilde{\mathscr{V}}$, let $\gamma = \gamma(v) = \sup_{1\leqslant i<\infty} R_i(v)$ and as in the definition of $\widetilde{\mathscr{V}}$, set

$$f_0 = 0,\ f_1 = 1,\ f_i = \prod_{k=1}^{i-1} v_k, \quad i \geqslant 2. \tag{9.5}$$

First, suppose that $\operatorname{supp}(v) = \{1, 2, \ldots, m-1\}$ for a finite m. Then $\operatorname{supp}(f) = \{1, 2, \ldots, m\}$ and

$$\frac{-\Omega f}{f}(i) = R_i(v) \leqslant \gamma, \qquad i = 1, 2, \ldots, m.$$

Hence,

$$\begin{aligned}
\gamma\sum_{k=1}^{m}\mu_k f_k^2 &\geqslant \sum_{k=1}^{m}\mu_k f_k(-\Omega f)(k)\\
&= \sum_{k=2}^{m+1}\mu_k a_k f_{k-1}(f_{k-1}-f_k) - \sum_{k=1}^{m}\mu_k a_k f_k(f_{k-1}-f_k) + \sum_{k=1}^{m}\mu_k c_k f_k^2\\
&= \sum_{k=1}^{m}\mu_k\big[a_k(f_{k-1}-f_k)^2 + c_k f_k^2\big] + \mu_{m+1}a_{m+1}f_m(f_m - f_{m+1})\\
&= \sum_{k=1}^{m+1}\mu_k\big[a_k(f_{k-1}-f_k)^2 + c_k f_k^2\big].
\end{aligned}$$

We have not only $\gamma \geqslant 0$ (actually $\gamma > 0$ when m is large enough since $c_i \not\equiv 0$) but also

$$\lambda_0 \leqslant \frac{D(f)}{\|f\|^2} \leqslant \gamma(v) \tag{9.6}$$

for all $v \in \widetilde{\mathscr{V}}$ with finite support.

(c) Next, we are going to prove (9.6) in the case that $v \in \widetilde{\mathscr{V}}$ with $v_i > 0$ for all $i \geqslant 1$. In this case, the positivity condition of v is not enough for the first equality in (9.2), as mentioned in Section 2 (above the proofs of Theorem 2.4 and Proposition 2.5). See also the specific situation given in the proof of Example 9.17 below. This explains why two additional conditions are included in the second union of the definition of $\widetilde{\mathscr{V}}$. The condition "$f \in L^2(\mu)$" is essential but not the one "$\Omega f/f \leqslant \eta$" since the eigenfunction g of λ_0 satisfies "$\Omega g/g = -\lambda_0$". To prove (9.6), without loss of generality, assume that $\gamma = \gamma(v) < \infty$. Otherwise, (9.6) is trivial. Clearly, $\gamma \geqslant R_1(v) = b_1(1 - v_1) + c_1 > -\infty$. Note that by assumptions, the function f possesses the following properties:

(i) $f > 0$ on E.
(ii) $f \in L^2(\mu)$ and then $P_t f \in L^2(\mu)$, where $P_t = (p_{ij}(t))$ is the minimal semigroup determined by the Dirichlet form.
(iii) $|\Omega f(i)| = \big|\sum_j q_{ij} f_j\big| \leqslant \max\{|\eta|, |\gamma|\} f_i$ for all $i \in E$.

Here, property (iii) comes from

$$-\eta f \leqslant -\Omega f \leqslant \gamma f.$$

Since $(p_{ij}(t))$ satisfies the forward Kolmogorov equation:

$$p_{ij}(t) = \delta_{ij} + \int_0^t \sum_k p_{ik}(s) q_{kj} \mathrm{d}s$$

and (i), it follows that

$$P_t f(i) = f_i + \sum_j \int_0^t \sum_k p_{ik}(s) q_{kj} f_j \mathrm{d}s.$$

By (ii), $P_t f(i) < \infty$ and is continuous in t. Because of this and (iii), the order of the last two sums and also the integration are exchangeable. This leads to

$$P_t f(i) \geqslant f_i - \gamma \int_0^t \sum_k p_{ik}(s) f_k \mathrm{d}s = f_i - \gamma \int_0^t P_s f(i) \mathrm{d}s, \qquad i \in E, \tag{9.7}$$

since by assumption $\Omega f \geqslant -\gamma f$. Therefore, we obtain

$$0 < D(f) = \lim_{t \downarrow 0} \frac{1}{t}(f, f - P_t f) \leqslant \lim_{t \downarrow 0} \frac{\gamma}{t} \int_0^t (f, P_s f) \mathrm{d}s = \gamma(v) \|f\|^2 < \infty.$$

Here, the first limit is due to (ii) and the first equality in (1.10), the last inequality comes from (i) and (9.7). We have thus proved that not only $\gamma > 0$ but also $f \in \mathscr{D}(D)$ and so we have returned to (9.6). In other words, (9.6) holds for all $v \in \widetilde{\mathscr{V}}$. By making infimum with respect to $v \in \widetilde{\mathscr{V}}$, we obtain

$$\lambda_0 \leqslant \inf_{v \in \widetilde{\mathscr{V}}} \sup_{i \in E} R_i(v).$$

(d) To prove the equality sign in the last formula holds, in view of proof (b), we have actually proved that for every finite m,

$$\begin{aligned}\lambda_0^{(m)} &:= \inf\{D(f) : f_0 = 0,\ f_i = 0 \text{ for all } i \geqslant m+1,\ \|f\| = 1\}\\ &\leqslant \inf_{v\in\widetilde{\mathscr{V}}_m}\ \sup_{1\leqslant i\leqslant m} R_i(v),\end{aligned} \tag{9.8}$$

where

$$\widetilde{\mathscr{V}}_m = \{v_i : v_i > 0 \text{ for } i < m,\ v_m = 0\}.$$

Actually, there is a $\bar v \in \widetilde{\mathscr{V}}_m$ such that $R_i(\bar v) = \lambda_0^{(m)} > 0$ for all $i\,(1 \leqslant i \leqslant m)$ since $m < \infty$ and then the equality sign in (9.8) holds. Therefore, the first equality in (9.2) holds since $\lambda_0^{(m)} \downarrow \lambda_0$ as $m \uparrow \infty$. $\square$

We now begin to study the estimate of λ_0. First, by Proposition 7.17, we have a simple upper bound:

$$\lambda_0 \leqslant \inf_{i\in E}(a_i + b_i + c_i).$$

Hence, $\lambda_0 = 0$ whenever $\lim_{n\to\infty}(a_n + b_n + c_n) = 0$. The next result provides us a finer upper bound. It is motivated from Theorem 3.1.

Proposition 9.4. *Let $\tilde c_i = c_i - \inf_i c_i$. Then*

$$\lambda_0 \leqslant \inf_{i\in E} c_i + \inf_{\ell\in E}\bigg(\sum_{i=1}^{\ell}\mu_i\bigg)^{-1}\inf_{E\ni m\geqslant \ell}\bigg[\bigg(\sum_{k=\ell}^{m}\frac{1}{\mu_k b_k}\bigg)^{-1} + \sum_{i=1}^{m}\mu_i\tilde c_i\bigg] \tag{9.9}$$

$$\leqslant \inf_{i\in E} c_i + \inf_{\ell\in E}\bigg(\sum_{i=1}^{\ell}\mu_i\bigg)^{-1}\bigg[\mu_\ell b_\ell + \sum_{i=1}^{\ell}\mu_i\tilde c_i\bigg]. \tag{9.10}$$

Proof. By Proposition 9.1, it is enough to consider the case that $\tilde c_i \equiv c_i$, i.e., $\inf_i c_i = 0$. Fix $\ell \leqslant m$ and define

$$\varphi_i = \varphi_i^{(\ell,m)} = \mathbb{1}_{\{i\leqslant m\}}\sum_{k=i\vee\ell}^{m}\frac{1}{\mu_k b_k}, \qquad i \in E.$$

Then

$$\mu\big(\varphi^2\big) = \sum_{i=1}^{\ell}\mu_i\varphi_\ell^2 + \sum_{i=\ell+1}^{m}\mu_i\varphi_i^2 \geqslant \varphi_\ell^2\sum_{i=1}^{\ell}\mu_i,$$

$$D(\varphi) = \sum_{k=\ell}^{m}\frac{1}{\mu_k b_k} + \varphi_\ell^2\sum_{i=1}^{\ell}\mu_i c_i + \sum_{i=\ell+1}^{m}\mu_i c_i\varphi_i^2 \leqslant \varphi_\ell + \varphi_\ell^2\sum_{i=1}^{m}\mu_i c_i.$$

Hence,

$$\frac{D(\varphi)}{\mu\big(\varphi^2\big)} \leqslant \bigg(\sum_{i=1}^{\ell}\mu_i\bigg)^{-1}\bigg[\varphi_\ell^{-1} + \sum_{i=1}^{m}\mu_i c_i\bigg].$$

Because $\varphi^{(\ell,m)} \in \mathscr{K}$, it follows that

$$\begin{aligned}
\lambda_0 &\leqslant \inf_{\ell\in E} \inf_{E\ni m\geqslant \ell} \frac{D(\varphi)}{\mu\big(\varphi^2\big)} \\
&\leqslant \inf_{\ell\in E} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \inf_{E\ni m\geqslant \ell} \bigg[\bigg(\sum_{k=\ell}^{m} \frac{1}{\mu_k b_k}\bigg)^{-1} + \sum_{i=1}^{m} \mu_i c_i\bigg] \\
&\leqslant \inf_{\ell\in E} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \inf_{E\ni m\geqslant \ell} \bigg[\mu_\ell b_\ell + \sum_{i=1}^{m} \mu_i c_i\bigg] \\
&= \inf_{\ell\in E} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \bigg[\mu_\ell b_\ell + \sum_{i=1}^{\ell} \mu_i c_i\bigg]. \quad \square
\end{aligned}$$

As an immediate consequence of (9.10), we obtain the following result.

Corollary 9.5. *If* $\sum_{i=1}^\infty \mu_i = \infty$ *and*

$$\varlimsup_{m\to\infty} \mu_m b_m \bigg(\sum_{i=1}^{m} \mu_i\bigg)^{-1} = 0,$$

then

$$\lambda_0 \leqslant \inf_{i\in E} c_i + \varliminf_{m\to\infty} \sum_{i=1}^{m} \mu_i \tilde{c}_i \Big/ \sum_{i=1}^{m} \mu_i \leqslant \inf_{i\in E} c_i + \varlimsup_{n\to\infty} \tilde{c}_n.$$

Proof. Without loss of generality, assume that $\tilde{c}_i \equiv c_i$.

By assumptions, it follows that

$$\begin{aligned}
&\varliminf_{\ell\to\infty} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \bigg[\mu_\ell b_\ell + \sum_{i=1}^{\ell} \mu_i c_i\bigg] \\
&\qquad\leqslant \varlimsup_{\ell\to\infty} \mu_\ell b_\ell \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} + \varliminf_{\ell\to\infty} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \sum_{i=1}^{\ell} \mu_i c_i \\
&\qquad= \varliminf_{\ell\to\infty} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \sum_{i=1}^{\ell} \mu_i c_i.
\end{aligned}$$

The first inequality now follows from (9.10).

To prove the second inequality, let $\gamma = \varlimsup_{n\to\infty} c_n \in [0,\infty]$. Then for every $\varepsilon > 0$, we have $\sup_{k\geqslant n} c_k \leqslant \gamma + \varepsilon$ for large enough n. Hence,

$$\sum_{i=1}^{\ell} \mu_i c_i = \sum_{i=1}^{n} \mu_i c_i + \sum_{i=n+1}^{\ell} \mu_i c_i \leqslant \sum_{i=1}^{n} \mu_i c_i + (\gamma+\varepsilon) \sum_{i=n+1}^{\ell} \mu_i, \qquad \ell > n.$$

We have thus obtained

$$\begin{aligned}
\varliminf_{\ell\to\infty} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \sum_{i=1}^{\ell} \mu_i c_i &\leqslant \varlimsup_{\ell\to\infty} \bigg(\sum_{i=1}^{\ell} \mu_i\bigg)^{-1} \bigg[\sum_{i=1}^{n} \mu_i c_i + (\gamma+\varepsilon) \sum_{i=n+1}^{\ell} \mu_i\bigg] \\
&= \gamma + \varepsilon
\end{aligned}$$

as required. $\square$

To study the lower estimate of λ_0, we observe that not every positive sequence (v_i) is useful for the lower estimate given in (9.2) since one may have $\inf_i R_i(v) < 0$. In order for $\inf_i R_i(v) \geqslant 0$, it is necessary that

$$0 < v_i \leqslant \frac{1}{b_i}\bigg(c_i + a_i + b_i - \frac{a_i}{v_{i-1}}\bigg).$$

From this, we obtain the following necessary condition:

$$\frac{a_{i+1}}{c_{i+1}+a_{i+1}+b_{i+1}} < v_i \leqslant x_i - \cfrac{y_i}{x_{i-1} - \cfrac{y_{i-1}}{x_{i-2} - \cfrac{y_{i-2}}{\ddots\, x_3 - \cfrac{y_3}{x_2 - \cfrac{y_2}{x_1}}}}},$$

where

$$x_i = \frac{c_i + a_i + b_i}{b_i}, \qquad y_i = \frac{a_i}{b_i}.$$

However, the condition is clearly not practical. Because of this reason, we are now going to introduce an alternative variational formula for the lower estimates.

For a given sequence (r_i), define an operator $I\!I^r = I\!I^{(r_i)}$ of "double sum" on the set of positive functions (f_i) as follows:

$$I\!I_1^r(f) = 0, \quad I\!I_i^r(f) = \sum_{k=1}^{i-1}\frac{1}{\mu_k b_k}\sum_{j=1}^{k} r_j\mu_j f_j = \sum_{j=1}^{i-1} r_j f_j \mu_j \sum_{k=j}^{i-1}\frac{1}{\mu_k b_k}, \quad E \ni i \geqslant 2.$$

Write $I\!I(f) = I\!I^{\mathbb{1}}(f)$. For a fixed sequence (c_i), let $\tilde{c}_i = c_i - \inf_i c_i$ and define

$$\mathscr{F} = \big\{f > 0 : f_i < f_1 + I\!I_i^{\tilde{c}}(f) \text{ for all } E \ni i \geqslant 2\big\}. \tag{9.11}$$

Clearly, if $\tilde{c}_1 > 0$, then every positive constant function belongs to $\mathscr{F}$. Otherwise, every $f > 0$ with $f_i < f_1$ for all $E \ni i \geqslant 2$ belongs to $\mathscr{F}$.

Theorem 9.6. *Let $I\!I^r$, $(\tilde{c}_i)$ and $\mathscr{F}$ be defined as above. Next, for each fixed $f \in \mathscr{F}$, define*

$$\xi = \xi_f = \begin{cases} \inf\limits_{E \ni i \geqslant 2} \dfrac{f_1 - f_i + I\!I_i^{\tilde{c}}(f)}{I\!I_i(f)}, & N = \infty, \\[2ex] \inf\limits_{E \ni i \geqslant 2} \dfrac{f_1 - f_i + I\!I_i^{\tilde{c}}(f)}{I\!I_i(f)} \bigwedge \dfrac{\sum_{j=1}^N \tilde{c}_j \mu_j f_j}{\sum_{j=1}^N \mu_j f_j}, & N < \infty, \end{cases} \tag{9.12}$$

$$\zeta(\eta, f) = \begin{cases} \inf\limits_{E \ni i \geqslant 2,\, \tilde{c}_i < \eta} \bigg[\tilde{c}_i + \dfrac{(\eta - \tilde{c}_i) f_i}{f_1 + I\!I_i^{\tilde{c}-\eta}(f)}\bigg], & \{E \ni i \geqslant 2 :\, \tilde{c}_i < \eta\} \neq \emptyset, \\ \eta, & \{E \ni i \geqslant 2 :\, \tilde{c}_i < \eta\} = \emptyset, \end{cases} \tag{9.13}$$

$$\eta \in [0, \xi].$$

Then we have

$$\lambda_0 \geqslant \inf_{i\in E} c_i + \zeta(\eta, f) \quad \textit{and} \quad \eta \geqslant \zeta(\eta, f), \qquad f \in \mathscr{F},\ \eta \in [0, \xi]. \tag{9.14}$$

Moreover, for fixed f, $\zeta(\eta, f)$ is increasing in η and furthermore,

$$\lambda_0 = \inf_{i\in E} c_i + \sup_{f\in\mathscr{F}} \zeta(\xi, f). \tag{9.15}$$

Remark 9.7. *To indicate the dependence on $(\tilde{c}_i)$, rewrite $\zeta(\eta, f)$ as $\zeta(\tilde{c}_i, \eta, f)$. Similarly, we have $\xi(\tilde{c}_i, f)$. Then for each $f \in \mathscr{F}$ and constant $\gamma \geqslant 0$, we have a shift property as follows:*

$$\xi(\tilde{c}_i + \gamma, f) = \gamma + \xi(\tilde{c}_i, f), \qquad \zeta(\tilde{c}_i + \gamma, \eta + \gamma, f) = \gamma + \zeta(\tilde{c}_i, \eta, f). \tag{9.16}$$

Hence, the use of $\inf_{i\in E} c_i$ in Theorem 9.6 is not essential but only for simplifying the computations. The same property holds for (9.10) *but not for* (9.9).

As will be illustrated later by Examples 9.17 and 9.19, it is not unusual that $\xi_f > \lambda_0$ for some $f \in \mathscr{F}$. In that case, we certainly have $\xi_f > \zeta(\xi_f, f)$. This means that ξ_f may not be a lower bound of λ_0 and so the use of $\zeta(\eta, f)$ in Theorem 9.6 is necessary.

Proof of Theorem 9.6. By Proposition 9.1, for simplicity, we assume that $\tilde{c}_i \equiv c_i$.

(a) First, we prove "$\lambda_0 \geqslant$" in (9.14). Fix $f \in \mathscr{F}$. Then $\xi = \xi_f \geqslant 0$. Without loss of generality, assume that $(\xi \geqslant)\, \eta > 0$. Otherwise, the assertion is trivial. Let

$$h_i = f_1 + I\!I_i^{c-\eta}(f), \qquad i \in E,\ \eta \in (0, \xi].$$

Since by (9.12),

$$f_1 - f_i + I\!I_i^{c}(f) \geqslant \eta I\!I_i(f) > 0$$

for $E \ni i \geqslant 2$ and $h_1 = f_1 > 0$, we have $h > 0$. Next, define $v_i = h_{i+1}/h_i$ ($v_0 := \infty$ and $v_N = 0$ if $N < \infty$). Then for $i : 2 \leqslant i < N$, since

$$h_i - h_{i+1} = I\!I_i^{c-\eta}(f) - I\!I_{i+1}^{c-\eta}(f) = \frac{1}{\mu_i b_i} \sum_{j=1}^{i} (\eta - c_j)\mu_j f_j,$$

we have

$$\begin{aligned}
a_i\big(1 - v_{i-1}^{-1}\big) &+ b_i(1 - v_i)\\
&= \frac{1}{h_i}\big[a_i(h_i - h_{i-1}) + b_i(h_i - h_{i+1})\big]\\
&= \frac{1}{h_i}\bigg[-\frac{a_i}{\mu_{i-1}b_{i-1}} \sum_{j=1}^{i-1}(\eta - c_j)\mu_j f_j + \frac{b_i}{\mu_i b_i}\sum_{j=1}^{i}(\eta - c_j)\mu_j f_j\bigg]\\
&= \frac{(\eta - c_i) f_i}{h_i}.
\end{aligned}$$

This also holds when $i=1$ (noting that $a_1=0$):

$$b_1(1-v_i)=\frac{b_1}{h_1}(h_1-h_2)=\frac{(\eta-c_1)f_1}{h_1}=\eta-c_1.$$

If $N<\infty$, then at $i=N$, by assumption

$$\eta\leqslant\xi\leqslant\sum_{j=1}^N c_j\mu_j f_j\bigg/\sum_{j=1}^N\mu_j f_j,$$

we get

$$a_N\big(1-v_{N-1}^{-1}\big)+b_N(1-v_N)=-\frac{a_N}{h_N\mu_{N-1}b_{N-1}}\sum_{j=1}^{N-1}(\eta-c_j)\mu_j f_j\geqslant\frac{(\eta-c_N)f_N}{h_N}.$$

Combining these facts together, we arrive at

$$R_i(v)=c_i+a_i\big(1-v_{i-1}^{-1}\big)+b_i(1-v_i)\geqslant c_i+\frac{(\eta-c_i)f_i}{h_i},\qquad i\in E.\tag{9.17}$$

We now show that the right-hand side of (9.17) is nonnegative for all i and so we have ruled out the useless case that $\inf_i R_i(v)<0$. Since $h>0$, the assertion is equivalent to

$$c_ih_i\geqslant(c_i-\eta)f_i,\qquad i\in E,$$

or

$$c_i\big[f_1-f_i+I\!I_i^c(f)\big]\geqslant\eta\big[c_iI\!I_i(f)-f_i\big].$$

This is trivial if $c_iI\!I_i(f)\leqslant f_i$ (in particular if $i=1$) since $f_1-f_i+I\!I_i^c(f)\geqslant 0$ for all $E\ni i\geqslant 2$ and $f\in\mathscr{F}$. Otherwise, by the definition of ξ and η, we have

$$f_1-f_i+I\!I_i^c(f)\geqslant\xi I\!I_i(f)\geqslant\eta I\!I_i(f)>\eta\big[I\!I_i(f)-f_i/c_i\big],\qquad E\ni i\geqslant 2.\tag{9.18}$$

We have thus proved the required assertion.

By Theorem 9.3 and (9.17), we obtain

$$\begin{aligned}\lambda_0&\geqslant\sup_{f\in\mathscr{F}}\inf_{i\in E}\left[c_i+\frac{(\eta-c_i)f_i}{f_1+I\!I_i^{c-\eta}(f)}\right]\\&=\sup_{f\in\mathscr{F}}\left\{\eta\wedge\inf_{E\ni i\geqslant 2}\left[c_i+\frac{(\eta-c_i)f_i}{f_1+I\!I_i^{c-\eta}(f)}\right]\right\}.\end{aligned}\tag{9.19}$$

Here, the last line is due to the fact that $I\!I_1^r(f)=0$.

(b) To prove the first assertion of the theorem, we show that for each i: $2\leqslant i\in E$,

$$c_i+\frac{(\eta-c_i)f_i}{f_1+I\!I_i^{c-\eta}(f)}\geqslant\eta\quad\text{iff}\quad c_i\geqslant\eta.$$

Clearly, the inequality is equivalent to

$$(\eta - c_i)f_i \geqslant (\eta - c_i)[f_1 + I\!I_i^{c-\eta}(f)].$$

The required assertion then follows since by (9.18), we already have

$$f_i \leqslant f_1 + I\!I_i^{c-\eta}(f).$$

As a consequence of the assertion, we have $\eta \geqslant \zeta(\eta, f)$. Now, from (9.19), it follows that

$$\lambda_0 \geqslant \sup_{f\in\mathscr{F}} \eta \wedge \zeta(\eta, f) = \sup_{f\in\mathscr{F}} \zeta(\eta, f).$$

This gives us the first assertion of the theorem.

(c) To prove the monotonicity of $\zeta(\eta, f)$ in η, let $\eta_1 < \eta_2 \leqslant \xi$. If $\{E \ni i \geqslant 2 : c_i < \eta_2\} = \emptyset$, then $\{E \ni i \geqslant 2 : c_i < \eta_1\} = \emptyset$ and so

$$\zeta(\eta_2, f) = \eta_2 > \eta_1 = \zeta(\eta_1, f).$$

If $\{E \ni i \geqslant 2 : c_i < \eta_1\} \neq \emptyset$, since $\{E \ni i \geqslant 2 : c_i < \eta_1\} \subset \{E \ni i \geqslant 2 : c_i < \eta_2\}$, we need only to show that

$$\frac{(\eta_2 - c_i)f_i}{f_1 + I\!I_i^{c-\eta_2}(f)} \geqslant \frac{(\eta_1 - c_i)f_i}{f_1 + I\!I_i^{c-\eta_1}(f)} \qquad \text{on } \{E \ni i \geqslant 2 : c_i < \eta_2\} \neq \emptyset.$$

Actually, this is enough even if $\{E \ni i \geqslant 2 : c_i < \eta_1\} = \emptyset$ in view of (b). Now, the required conclusion is trivial on the set $\{E \ni i \geqslant 2 : \eta_1 \leqslant c_i < \eta_2\}$. Hence, it suffices to show that

$$\frac{\eta_2 - c_i}{f_1 + I\!I_i^{c-\eta_2}(f)} \geqslant \frac{\eta_1 - c_i}{f_1 + I\!I_i^{c-\eta_1}(f)} \qquad \text{on } \{E \ni i \geqslant 2 : c_i < \eta_1\}.$$

A simple computation shows that this is equivalent to

$$f_1 + I\!I_i^{c}(f) \geqslant c_i I\!I_i(f),$$

which holds on $\{E \ni i \geqslant 2 : c_i < \eta_1\}$ in view of (9.12) and $\xi > \eta_1$.

(d) To prove (9.15), it suffices to show that the equality in (9.19) holds for $\eta = \xi$. Noting that the right-hand side of (9.19) is nonnegative, without loss of generality, we may assume that $\lambda_0 > 0$. Then, by Proposition 2.1, the eigenfunction $g > 0$ of λ_0 satisfies

$$\mu_k b_k (g_k - g_{k+1}) = \sum_{j=1}^{k} (\lambda_0 - c_j)\mu_j g_j, \qquad k \in E,\ g_{N+1} = 0 \text{ if } N < \infty.$$

Hence,

$$g_1 - g_i = I\!I_i^{\lambda_0 - c}(g), \quad i \in E, \qquad \sum_{j=1}^{N} (\lambda_0 - c_j)\mu_j g_j = 0 \quad \text{if } N < \infty$$

and furthermore, $g \in \mathscr{F}$. It follows that

$$\frac{g_1 - g_i + II_i^c(g)}{II_i(g)} \equiv \lambda_0, \qquad E \ni i \geqslant 2,$$

$$\frac{\sum_{j=1}^N c_j \mu_j g_j}{\sum_{j=1}^N \mu_j g_j} = \lambda_0 \qquad \text{if } N < \infty,$$

$$c_i + \frac{(\lambda_0 - c_i) g_i}{g_1 + II_i^{c-\lambda_0}(g)} \equiv \lambda_0, \qquad i \in E.$$

Therefore, $\xi_g = \lambda_0$, and furthermore, the equality sign in (9.19) is attained at $(f, \eta) = (g, \lambda_0)$. $\square$

We now make a rough comparison of Theorems 9.6 and 9.3 for the lower estimate. See also the comment below the proof of Corollary 9.9.

Remark 9.8. *For a given positive sequence (v_i) such that $\inf_{i \in E} R_i(v) := \gamma_v \geqslant 0$, corresponding to the sequence (f_i) and ξ_f defined by (9.5) and (9.12), respectively, we have $\xi_f \geqslant \gamma_v$.*

Proof. From the assumption

$$R_i(v) = c_i + a_i\big(1 - v_{i-1}^{-1}\big) + b_i(1 - v_i) \geqslant \gamma_v =: \gamma, \qquad i \in E,$$

it follows that

$$f_k - f_{k+1} \geqslant \frac{1}{\mu_k b_k} \sum_{j=1}^k (\gamma - c_j) \mu_j f_j,$$

and then

$$f_1 - f_i \geqslant \sum_{k=1}^{i-1} \frac{1}{\mu_k b_k} \sum_{j=1}^k (\gamma - c_j) \mu_j f_j = II_i^{\gamma - c}(f), \qquad i \in E.$$

To prove our assertion, without loss of generality, assume that $\gamma > 0$. Then it is clear not only that $f \in \mathscr{F}$ but also $\xi_f \geqslant \gamma$. $\square$

As a complement to Remark 9.8, it would be nice if we could show that

$$c_i + \frac{(\xi_f - c_i) f_i}{f_1 + II_i^{c-\xi_f}(f)} \geqslant \gamma_v \quad \text{on the set } \{E \ni i \geqslant 2 : c_i < \xi_f\}.$$

This holds obviously on the subset $\{\gamma_v \leqslant c_i < \xi_f\}$, but is not clear on the subset $\{E \ni i \geqslant 2 : c_i < \gamma_v\}$.

The next result is a particular application of Theorem 9.6. It is a complement of Corollary 9.5. The combination of Proposition 9.4 and Corollary 9.5 with Corollary 9.9 below indicates that when $\lambda_0(a_i, b_i, 0) = 0$, the condition $\lim_{n \to \infty} c_n > 0$ is crucial for $\lambda_0(a_i, b_i, c_i) > 0$. This is more or less clear in terms of the Feynman-Kac formula:

$$P_t^c f(x) = \mathbb{E}^x \Big[f(X_t) e^{-\int_0^{t \wedge \tau} c_{X_s} \mathrm{d}s} \Big],$$

where $\{P_t^c\}_{t \geqslant 0}$ is the minimal semigroup generalized by the operator with rates (a_i, b_i, c_i), $\{X_t\}_{0 \leqslant t < \tau}$ is the minimal process with rates (a_i, b_i), and τ is the life time of $\{X_t\}$. Note that $\lambda_0(a_i, b_i, 0) > 0$, and hence, $\lambda_0(a_i, b_i, c_i) > 0$ if the uniqueness condition (1.2) fails. Otherwise, $\tau = \infty$.

Corollary 9.9. *Let $\varepsilon \in (0,1)$. Define*

$$\xi_\varepsilon = \begin{cases} \inf\limits_{E\ni i\geqslant 2} \dfrac{1-\varepsilon+\tilde{c}_1 z_i+\varepsilon x_i}{z_i+\varepsilon y_i}, & N=\infty, \\ \inf\limits_{E\ni i\geqslant 2} \dfrac{1-\varepsilon+\tilde{c}_1 z_i+\varepsilon x_i}{z_i+\varepsilon y_i} \bigwedge \dfrac{\tilde{c}_1+\varepsilon\sum_{j=2}^N \tilde{c}_j\mu_j}{1+\varepsilon\sum_{j=2}^N \mu_j}, & N<\infty, \end{cases} \tag{9.20}$$

$$\zeta_\varepsilon = \begin{cases} \inf\limits_{E\ni i\geqslant 2:\, \tilde{c}_i<\xi_\varepsilon} \left[\tilde{c}_i + \dfrac{\varepsilon(\xi_\varepsilon-\tilde{c}_i)}{1+\tilde{c}_1 z_i+\varepsilon x_i-\xi_\varepsilon(z_i+\varepsilon y_i)}\right], & \\ \qquad\qquad\qquad\qquad \{E\ni i\geqslant 2:\, \tilde{c}_i<\xi_\varepsilon\}\neq\emptyset, & \\ \xi_\varepsilon, \qquad\qquad\qquad\quad \{E\ni i\geqslant 2:\, \tilde{c}_i<\xi_\varepsilon\}=\emptyset, & \end{cases} \tag{9.21}$$

where

$$x_i = \sum_{2\leqslant j\leqslant i-1} \tilde{c}_j\,\mu_j\,\nu[j,i-1], \qquad y_i = \sum_{2\leqslant j\leqslant i-1} \mu_j\,\nu[j,i-1], \qquad z_i = \nu[1,i-1],$$

and $\nu[i,j] = \sum_{i\leqslant k\leqslant j}(\mu_k b_k)^{-1}$. Then we have $\lambda_0 \geqslant \inf_{i\in E} c_i + \sup_{\varepsilon\in(0,1)}\zeta_\varepsilon$. The same conclusion holds if ξ_ε in (9.21) is replaced by $\eta\in[0,\xi_\varepsilon]$. In particular, if $\varliminf_{n\to\infty} c_n > 0$, then $\lambda_0>0$.

Proof. (a) The main assertion of the corollary is an application of Theorem 9.6 to the specific $f\in\mathscr{F}$: $f_1=1$, $f_i=\varepsilon\in(0,1)\,(E\ni i\geqslant 2)$, for which we have

$$I\!I_1^r(f)=0, \quad I\!I_i^r(f) = r_1\sum_{1\leqslant k\leqslant i-1}\frac{1}{\mu_k b_k} + \varepsilon\sum_{2\leqslant j\leqslant i-1} r_j\mu_j \sum_{j\leqslant k\leqslant i-1}\frac{1}{\mu_k b_k}, \quad E\ni i\geqslant 2.$$

Then (9.20) and (9.21) follows from (9.12) and (9.13), respectively.

We now prove the particular assertion for which $N=\infty$.

(b) If (1.2) does not hold, then $\lambda_0(a_i,b_i,0)>0$ by Theorem 3.1, and so $\lambda_0>0$ by part (3) of Proposition 9.1. Similarly, if $\inf_i c_i>0$, then we have again $\lambda_0>0$. Thus, without loss of generality, assume that

$$\inf_i c_i = 0 \text{ and (1.2) holds.}$$

(c) With the test function f given in (a), by (9.12), we have

$$\xi_\varepsilon = \inf_{i\geqslant 2}\frac{1-\varepsilon+I\!I_i^c(f)}{I\!I_i(f)}.$$

By assumption, there exist $\gamma>0$ and $m\geqslant 2$ such that $c_i>\gamma$ for all $i\geqslant m$. Certainly, we have

$$\xi_\varepsilon \geqslant \inf_{2\leqslant i\leqslant m}\frac{1-\varepsilon+I\!I_i^c(f)}{I\!I_i(f)} \bigwedge \inf_{i>m}\frac{I\!I_i^c(f)}{I\!I_i(f)}.$$

For $i > m$, we have

$$\frac{I\!I_i^c(f)}{I\!I_i(f)} \geqslant \sum_{j=m}^{i-1} c_j f_j \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k} \Big/ \sum_{j=1}^{i-1} f_j \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k}$$
$$> \varepsilon\gamma \sum_{j=m}^{i-1} \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k} \Big/ \sum_{j=1}^{i-1} \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k}.$$

By assumption (1.2), the right-hand side goes to $\varepsilon\gamma > 0$ as $i \to \infty$. It follows that $\inf_{i>m} I\!I_i^c(f)/I\!I_i(f) > 0$, and furthermore, there exists $\eta \in (0, \gamma)$ such that $\xi_\varepsilon > \eta$.

(d) Noting that the set $\{i \geqslant 2 : c_i < \eta\} \subset \{i : 2 \leqslant i < m\}$ is finite, by (9.13), we have

$$\zeta(\eta, f) = \inf_{i \geqslant 2,\, c_i < \eta} \left[c_i + \frac{(\eta - c_i) f_i}{f_1 + I\!I_i^{c-\eta}(f)} \right] \geqslant \min_{i \geqslant 2,\, c_i < \eta} \frac{(\eta - c_i) f_i}{f_1 + I\!I_i^{c-\eta}(f)} > 0.$$

Now, by Theorem 9.6 or proof (a) above, we conclude that $\lambda_0 > 0$. $\square$

From proof (c) above, we have seen that when $N = \infty$,

$$\xi_\varepsilon > 0 \quad \text{iff} \quad \inf_{i>m} I\!I_i^{\tilde{c}}(\mathbb{1})/I\!I_i(\mathbb{1}) > 0 \quad \text{for all } m \geqslant 2. \tag{9.22}$$

Note that

$$\inf_{i>m} \frac{I\!I_i^{\tilde{c}}(\mathbb{1})}{I\!I_i(\mathbb{1})} \geqslant \inf_{i \geqslant 1} \sum_{j=1}^{i} \mu_j \tilde{c}_j \Big/ \sum_{j=1}^{i} \mu_j$$

and the right-hand side is positive iff

$$\varliminf_{m\to\infty} \sum_{j=1}^{m} \mu_j \tilde{c}_j \Big/ \sum_{j=1}^{m} \mu_j > 0. \tag{9.23}$$

Thus, Corollary 9.9 is qualitatively consistent with Corollary 9.5.

In view of Remark 9.8, it is not obvious that Theorem 9.6 improves Theorem 9.3. An easier way to see the improvement is as follows. Recall that the last assertion of Corollary 9.9 is deduced in terms of the test function f used in its proof (a). For which, the corresponding sequence (v_i) is $v_1 = 1/2$ and $v_i = 1$ for all $i \geqslant 2$. Inserting this into $R(v)$, we get

$$\inf_{i \geqslant 1} R_i(v) = (c_1 + b_1/2) \wedge (c_2 - a_2) \wedge \inf_{i \geqslant 3} c_i.$$

Thus, for $\inf_{i \geqslant 1} R_i(v) > 0$, it is necessary that $\inf_{i \geqslant 3} c_i > 0$, which is clearly much stronger than the last condition $\varliminf_{n\to\infty} c_n > 0$ used in Corollary 9.9.

As Proposition 9.4, the next result is also motivated from Theorem 3.1.

Corollary 9.10. *An explicit lower estimate can be obtained by Theorem 9.6 using the specific test function* $f^{(m)}$*:*

$$f_i^{(m)} = \left(\sum_{j=i\wedge m}^{m} \frac{1}{\mu_j b_j}\right)^{1/2}, \qquad i \in E,$$

where m *may be optimized over* $\{m \in E : m \geqslant 2\}$ *(or over* E *if* $\tilde{c}_1 > 0$*).*

We now show that some special killing (or Schrödinger) case can be regarded as a perturbation of the one without killing. To do so, fix constants $\beta, \gamma > 0$, and define

$$\begin{gathered}
\hat{a}_i = b_{i-1}, \quad 2 \leqslant i < N+1, \qquad \hat{b}_i = a_{i+1}, \quad 1 \leqslant i < N, \\
\hat{a}_1 = \beta, \quad \hat{b}_N = \gamma \quad \text{if } N < \infty; \\
\check{a}_i = a_{i+1}, \quad 0 \leqslant i < N, \qquad \check{b}_i = b_i, \quad 1 \leqslant i < N+1. \\
\check{b}_0 = \beta, \quad \check{a}_N = \gamma \quad \text{if } N < \infty.
\end{gathered} \tag{9.24}$$

Note that $(\hat{a}_i, \hat{b}_i)$ and $(\check{a}_i, \check{b}_i)$ are dual each other in the sense of Section 5 but they are clearly different from (a_i, b_i). Recall that $a_1 = 0$ and $b_N = 0$ by convention. Next, let (c_i) satisfy

$$c_i \geqslant \begin{cases} a_{i+1} - a_i - b_i + b_{i-1}, & 2 \leqslant i < N, \\ a_2 - b_1 + \beta, & i = 1, \\ \gamma - a_N + b_{N-1}, & i = N < \infty. \end{cases} \tag{9.25}$$

Note that the right-hand side of (9.25) can be negative. Conversely, for given rates $(\hat{a}_i, \hat{b}_i)$, the inverse transform is as follows:

$$\begin{gathered}
a_i = \hat{b}_{i-1}, \quad 2 \leqslant i < N+1, \qquad b_i = \hat{a}_{i+1}, \quad 1 \leqslant i < N, \\
c_i \geqslant \hat{b}_i - \hat{b}_{i-1} - \hat{a}_{i+1} + \hat{a}_i, \qquad 1 \leqslant i < N+1 \ \ (\text{or } i \in E).
\end{gathered} \tag{9.26}$$

Proposition 9.11. *Suppose that the given rates* $(a_i, b_i, c_i : i \in E)$ *satisfy (9.25). Define* $\lambda_0(a_i, b_i, c_i)$ *as in Proposition 9.1 without preassuming that* $c_i \geqslant 0$ *for all* $i \in E$*. Next, define* $(\hat{a}_i, \hat{b}_i)$ *and* $(\check{a}_i, \check{b}_i)$ *by (9.24).*

(1) *If* $\sum_{i=2}^N \mu_i b_{i-1}^{-1} = \infty$*, then* $\lambda_0(a_i, b_i, c_i) \geqslant \hat{\lambda}_0$*, where* $\hat{\lambda}_0$ *is defined by (4.1) with rates* $(\hat{a}_i, \hat{b}_i)$*.*

(2) *Otherwise,* $\lambda_0(a_i, b_i, c_i) \geqslant \check{\lambda}_1$*, where* $\check{\lambda}_1$ *is defined by (6.1) with rates* $(\check{a}_i, \check{b}_i)$*.*

(3) *The equality sign of the conclusions in parts* (1) *and* (2) *holds provided it does in* (9.25)*.*

Proof. (a) As an application of Proposition 9.1, without loss of generality, we may and will assume that the equality sign for c_i in (9.26) holds. Then, we prove that the equality sign of the conclusions in parts (1) and (2) holds.

Clearly, we have

$$\hat\mu_1 = 1,\ \ \hat\mu_1 \hat a_1 = \beta,\ \ \hat\mu_i = \mu_i^{-1},\ \ \hat\mu_i \hat a_i = \frac{b_{i-1}}{\mu_i}, \qquad 2 \leqslant i < N+1. \tag{9.27}$$

(b) Recall the operators:

$$\Omega f(i) = b_i(f_{i+1} - f_i) + a_i(f_{i-1} - f_i) - c_i f_i,$$
$$\widehat\Omega f(i) = \hat b_i(f_{i+1} - f_i) + \hat a_i(f_{i-1} - f_i), \quad f \in \mathscr{K}, f_0 = 0, f_{N+1} = 0 \text{ if } N < \infty.$$

Clearly, $\lambda_0(a_i, b_i, c_i)$ is the principal eigenvalue of Ω and the idea is describing it in terms of the first eigenvalue $\hat\lambda_{\min}$ of $\widehat\Omega$. Let U be the diagonal matrix with diagonal elements $(\mu_i : i \in E)$. Then U^{-1} is simply the diagonal matrix with diagonal elements $(\hat\mu_i : i \in E)$. For each function h with $h_0 = 0$ and $h_{N+1} = 0$ if $N < \infty$, by (9.27), (9.24) and (9.26), we have

$$\begin{aligned}
\big(\Omega U^{-1} h\big)(i) &= b_i\big(\hat\mu_{i+1} h_{i+1} - \hat\mu_i h_i\big) + a_i\big(\hat\mu_{i-1} h_{i-1} - \hat\mu_i h_i\big) - c_i \hat\mu_i h_i \\
&= \hat a_{i+1}\big(\hat\mu_{i+1} h_{i+1} - \hat\mu_i h_i\big) + \hat b_{i-1}\big(\hat\mu_{i-1} h_{i-1} - \hat\mu_i h_i\big) \\
&\quad - \big(\hat b_i - \hat b_{i-1} - \hat a_{i+1} + \hat a_i\big)\hat\mu_i h_i \\
&= \big(\hat a_{i+1}\hat\mu_{i+1} h_{i+1} - \hat b_i \hat\mu_i h_i\big) + \big(\hat b_{i-1}\hat\mu_{i-1} h_{i-1} - \hat a_i \hat\mu_i h_i\big) \\
&= \hat\mu_i \hat b_i (h_{i+1} - h_i) + \hat\mu_i \hat a_i (h_{i-1} - h_i) \\
&= \hat\mu_i \widehat\Omega\, h(i) \\
&= \big(U^{-1}\widehat\Omega\, h\big)(i), \qquad 2 \leqslant i < N.
\end{aligned}$$

It is easy to check that the identity holds also for $i = 1$ and $i = N$, and then for all $i \in E$. Multiplying U from the left on the both sides, we obtain

$$U \Omega U^{-1} = \widehat\Omega. \tag{9.28}$$

Furthermore, we get

$$\langle f, \Omega g\rangle_\mu = \langle U^{-1} f, (U\Omega U^{-1}) U g\rangle_\mu = \langle U f, (U\Omega U^{-1}) U g\rangle_{\hat\mu} = \langle \hat f, \widehat\Omega \hat g\rangle_{\hat\mu}$$

for all $f, g \in \mathscr{K}$, where the mapping $f \to \hat f := Uf$ is an isometry from $L^2(\mu)$ to $L^2(\hat\mu)$. Since $f \in \mathscr{K}$ iff $\hat f \in \mathscr{K}$, it follows that the operators Ω and $\widehat\Omega$ with the same core $\mathscr{K}$ are isospectral. In particular, $\lambda_0(a_i, b_i, c_i) = \hat\lambda_{\min}$.

(c) For assertion (1), since $\sum_i \big(\hat\mu_i \hat b_i\big)^{-1} = \infty$ by assumption, it follows that $N = \infty$ and the Dirichlet form corresponding to $\widehat\Omega$ is regular by Proposition 1.3. Hence, the minimal and the maximal domains of the Dirichlet form are coincided. Therefore, $\hat\lambda_{\min}$ is equal to $\lambda_0^{(4.1)}$ replacing the original rates (a_i, b_i) by $\big(\hat a_i, \hat b_i\big)$.

For assertion (2), since $\hat a_1 > 0$ and $\hat b_N > 0$, we come to the setup of Section 7: $\hat\lambda_{\min} = \lambda_0^{(7.1)}$ with (a_i, b_i) replaced by $\big(\hat a_i, \hat b_i\big)$. Next, because of $\sum_i \big(\hat\mu_i \hat b_i\big)^{-1} < \infty$, by Theorem 7.1, it turns out $\hat\lambda_{\min} = \check\lambda_1$ in terms of the dual rates $\big(\check a_i, \check b_i\big)$ of $\big(\hat a_i, \hat b_i\big)$. $\square$

We now summarize our main qualitative result about λ_0. The three parts given below are obtained by Corollary 9.9, Proposition 9.1 plus Proposition 1.3, and Corollary 9.5, respectively.

Summary 9.12. *We have $\lambda_0 > 0$ whenever $N < \infty$. Next, let $N = \infty$. Then*

(1) $\lambda_0 > 0$ *if* $\varliminf_{n\to\infty} c_n > 0$ *(in particular, if* $\inf_i c_i > 0$*).*

(2) $\lambda_0 = \lambda_0(a_i, b_i, c_i) > 0$ *if* $\lambda_0\big(a_i, b_i, c_i \mathbb{1}_{\{1\}}\big) > 0$ *which can be checked case by case by*

  (i) Theorem 3.1 when $c_1 = 0$ *and*

$$\sum_{i\geqslant 1} \frac{1}{\mu_i a_i} < \infty; \tag{9.29}$$

  (ii) Theorems 7.1 and 6.2 when $c_1 > 0$ *and (9.29) holds;*

  (iii) Theorem 4.2 when $c_1 > 0$ *but (9.29) fails.*

(3) $\lambda_0 = 0$ *if*

$$\varliminf_{m\to\infty} \sum_{i=1}^m \mu_i c_i \Big/ \sum_{k=1}^m \mu_k = 0, \quad \varlimsup_{m\to\infty} \mu_m b_m \bigg(\sum_{i=1}^m \mu_i\bigg)^{-1} = 0 \quad \text{and} \quad \sum_i \mu_i = \infty.$$

Open problem 9.13 (Explicit criterion for $\lambda_0 > 0$). As will be seen soon in Example 9.18 below, for $\lambda_0 > 0$, it can happen that $\varliminf_{n\to\infty} c_n = 0$. Hence, the simple condition "$\varliminf_{n\to\infty} c_n > 0$" in part (1) is sufficient only but not necessary. Naturally, this condition becomes necessary for the first one in part (3) for which a sufficient condition is $\lim_{n\to\infty} c_n = 0$. Thus, it is more or less satisfactory whenever (c_n) has a limit. Otherwise, there is a gap. In contrast with the first condition in part (3), condition (9.23) is sufficient for $\xi_\varepsilon > 0$ but there is still a distance to deduce the positivity of λ_0 in view of Corollary 9.9.

Next, since we are dealing with the minimal Dirichlet form, a general criterion for Hardy-type inequalities (cf. [12; Theorems 7.1 and 7.2]) which was successfully used in Section 8, is also available in the present situation, hence, there is already a criterion for $\lambda_0 > 0$ in terms of capacity which is unfortunately not explicit. More seriously, the technique to produce an explicit result used in [12; pages 134–136] does not work at the beginning (replacing a finite number of disjointed finite intervals $\{K_i\}$ by the connected one $[\min \cup_i K_i, \max \cup_i K_i]$) in the present setup. Thus, it is still an unsolved problem to figure out an explicit criterion for $\lambda_0 > 0$ in the present setup.

It is our position to illustrate by examples the application of the results obtained in this section. First, by using Proposition 9.11 and (9.26), it is easy to transfer the examples given in Sections 3 and 6 to the present context. However, most of the resulting killing rates are rather simple. We are now going to construct some new examples, all of them are out of the scope of Proposition 9.11. In the most cases, we use simple (a_i, b_i) and pay more attention on (c_i). Let us begin with the following simplest case.

Example 9.14. *Let*

$$Q = \begin{pmatrix} -b_1 - c_1 & b_1 \\ a_2 & -a_2 - c_2 \end{pmatrix}.$$

Then as in Examples 7.5 (2), we have

$$\lambda_0 = \frac{1}{2}\Big(a_2 + b_1 + c_1 + c_2 - \sqrt{(a_2 + c_2 - b_1 - c_1)^2 + 4a_2 b_1}\,\Big),$$

with eigenvector

$$g = \bigg(\frac{1}{2a_2}\Big[a_2 + c_2 - b_1 - c_1 + \sqrt{(a_2 + c_2 - b_1 - c_1)^2 + 4a_2 b_1}\,\Big],\ 1\bigg).$$

Even in such a simple case, the role for λ_0 played by the parameters a_i, b_i, and c_i is ambiguous. For instance, since $\tilde{c}_1 - \tilde{c}_2 = c_1 - c_2$ ($\tilde{c}_k := c_k - c_1 \wedge c_2$), one can separate out the constant $c_1 \wedge c_2$ from the above expression of λ_0. However, this obvious separation property becomes completely mazed for the next example having three states only.

Example 9.15. *Let*

$$Q = \begin{pmatrix} -b_1 - c_1 & b_1 & 0 \\ a_2 & -a_2 - b_2 - c_2 & b_2 \\ 0 & a_3 & -a_3 - c_3 \end{pmatrix}.$$

Then

$$\lambda_0 = -\frac{1}{3}\gamma_1 + 2\sqrt{\frac{-U}{3}}\,\cos\bigg[\frac{1}{3}\,\textit{arc cos}\bigg(-\frac{V}{2}\Big(\frac{-U}{3}\Big)^{-3/2}\bigg) + \frac{2\pi}{3}\bigg], \tag{9.30}$$

with eigenvector

$$g = \bigg\{\frac{b_1(a_3 + c_3 - \lambda_0)}{a_3(b_1 + c_1 - \lambda_0)},\ 1 + \frac{c_3 - \lambda_0}{a_3},\ 1\bigg\},$$

where

$$U = \gamma_2 - \gamma_1^2/3, \qquad V = \gamma_3 - \gamma_1\gamma_2/3 + 2(\gamma_1/3)^3,$$

and $\lambda^3 + \gamma_1\lambda^2 + \gamma_2\lambda + \gamma_3$ *the eigenpolynomial of* $-Q$ *with coefficients:*

$$\begin{aligned}
\gamma_1 &= -a_2 - a_3 - b_1 - b_2 - c_1 - c_2 - c_3,\\
\gamma_2 &= a_2a_3 + b_1a_3 + c_1a_3 + c_2a_3 + b_1b_2 + a_2c_1 + b_2c_1 + b_1c_2 + c_1c_2 + a_2c_3\\
&\quad + b_1c_3 + b_2c_3 + c_1c_3 + c_2c_3,\\
\gamma_3 &= -a_2a_3c_1 - a_3b_1c_2 - a_3c_1c_2 - b_1b_2c_3 - a_2c_1c_3 - b_2c_1c_3 - b_1c_2c_3 - c_1c_2c_3.
\end{aligned}$$

Proof. Since the eigenvalues of $-Q$ are all real, it is easier to write them down. By using the notation given above, the eigenvalues of $-Q$ can be expressed as

$$-\frac{1}{3}\gamma_1 + 2\sqrt{\frac{-U}{3}}\,\cos\bigg[\frac{1}{3}\,\mathrm{arc\,cos}\bigg(-\frac{V}{2}\Big(\frac{-U}{3}\Big)^{-3/2}\bigg) + \frac{2k\pi}{3}\bigg], \qquad k = 0, 1, 2.$$

Among them, the minimal one is λ_0 given in (9.30). Clearly, the solution is indeed rather complicated in view of the coefficients of the eigenpolynomial. $\square$

To see the role played by the killing rate (c_i), in the following examples, we restrict ourselves to the case that $\lambda_0(a_i, b_i, 0) = 0$ (and then $N = \infty$). The examples are arranged according the increasing order of the polynomial rates (a_i) and (b_i). Actually, all the examples in the paper are either standard or constructed by using simple rates and simple eigenfunctions. They are used first as a guidance of the study and then to justify the power of the theoretic results.

In contract to the explosive case (cf. Theorem 3.1), λ_0 can still be zero for the process having positive killing rate, as shown by the following example.

Example 9.16. *Let $b_1 = 1$, $a_i = b_i = 1$ for $i \geqslant 2$, and let (c_i) satisfy $\lim_{n\to\infty} c_n = 0$. Then we have $\lambda_0 = 0$, even though c_i can be very large locally.*

Proof. Apply Corollary 9.5. $\square$

Example 9.17. *Let $a_i = b_i = 1$ for $i \geqslant 2$ and $c_i = \beta^{-1}(\beta-1)^2$ $(\beta > 0)$ for $i \geqslant 1$. Then for every $a_1 \geqslant 0$ and $b_1 > 0$, we have $\lambda_0 = \beta^{-1}(\beta-1)^2$.*

Proof. Since $\lambda_0(a_i, b_i, 0) = 0$ and (c_i) is a constant, this is a consequence of part (3) of Proposition 9.1.

Note that the lower estimates given by Proposition 9.1, Theorem 9.3, and (9.4) are all sharp for this example. To see this, simply choose $v_i \equiv 1$ in (9.2) and (9.4). We now consider a more specific situation: $a_1 = 0$, $b_1 = 1-\beta$ and $\beta \in (0,1)$. If we set $\bar{v}_i \equiv \beta^{-1}$, then it is easy to check that $R_i(\bar{v}) \equiv 0$ and so $\inf_{v>0}\sup_{i\geqslant 1} R_i(v) = 0$. This shows that the truncating procedure used in Theorem 9.3 for the upper estimate is necessary in the case that the function f defined by (9.5) does not belong to $L^2(\mu)$, even though $R_i(v)$ is a constant. In the present case, $\bar{v}_i > 1$ for all i and so the corresponding function f is strictly increasing. Since μ_i is a constant for $i \geqslant 2$, it is clear that $\sum_i \mu_i = \infty$ and then $f \notin L^2(\mu)$. $\square$

Example 9.18. *Let $a_1 = 0$, $b_1 = 5/2$, $a_i = 2$ and $b_i = 1$ for $i \geqslant 2$, $c_i = 0$ for odd i and $c_i = 13/6$ for even i. Then $\lambda_0 = 5/6$. The upper bound provided by (9.9) is approximately* 1.03*. For the lower estimate, Proposition 9.11 is available but not Corollary 9.9.*

Proof. (a) Let $v_i \equiv 1 + (-1)^i/3$. Then it is easy to check that $R_i(v) \equiv 5/6$. Next, define

$$g_1 = 1, \quad g_n = \prod_{k=1}^{n-1} v_k, \qquad n \geqslant 2. \tag{9.31}$$

We claim that $g \in L^2(\mu)$ by using Kummer's test. To do so, note that to study the convergence/divergence of the series $\sum_n \mu_n g_n^2$, the constant κ defined by (3.13) takes a simpler form as follows:

$$\kappa = \lim_{n\to\infty} n\left(\frac{a_{n+1}}{b_n v_n^2} - 1\right). \tag{9.32}$$

Now, because $g \in L^2(\mu)$ and

$$-\Omega g/g = R(v) = 5/6,$$

we have $\lambda_0 = 5/6$ by Theorem 9.3. Clearly, this eigenfunction g of λ_0 is not monotone since $g_{i+1}/g_i = v_i = 2/3$ for odd i and $= 4/3$ for even i.

(b) Next, we study the upper estimates of λ_0. First, we have

$$\mu_1 = 1, \ \ \mu_i = \frac{5}{2^i}, \quad i \geqslant 2; \qquad \mu_i b_i = \frac{5}{2^i}, \quad i \geqslant 1.$$

The upper bound provided by (9.9) is approximately 1.03.

(c) For a lower estimate, we apply Proposition 9.11 (2). The modified rates are as follows: $\breve{a}_i \equiv 2\,(i \geqslant 1)$, $\breve{b}_1 = 5/2$, and $\breve{b}_i \equiv 1\,(i \geqslant 2)$. However, (c_i) does not satisfy (9.25) at $i = 2$. We now replace (c_i) by $(\tilde{c}_i := c_i + 1/6)$ and choose $\breve{b}_0 = \beta = 2/3$. Then $(\tilde{c}_i)$ satisfy (9.25). With the modified $(\tilde{c}_i)$, we are in the ergodic case, and moreover, $\breve{\lambda}_1 = \big(\sqrt{2} - 1\big)^2$ with eigenfunction $\breve{g}$: $\breve{g}_0 = -1$ and

$$\breve{g}_i = \frac{1}{20} 2^{i/2}\Big[-101 + 60\sqrt{2} + \Big(41 - 25\sqrt{2}\Big)\, i\Big], \qquad i \geqslant 1.$$

Therefore, by Proposition 9.11 (2), we obtain $\lambda_0(a_i, b_i, \tilde{c}_i) \geqslant \big(\sqrt{2}-1\big)^2$. Returning to the original (c_i) by Proposition 9.1 (2), we get a rough lower bound as follows:

$$\lambda_0 = \lambda_0(a_i, b_i, \hat{c}_i) - \frac{1}{6} \geqslant \frac{17}{6} - 2\sqrt{2} \approx 0.005.$$

Before moving further, let us remark that if only $\breve{b}_0$ is changed from 2/3 to 1/2, then for the $\big(\breve{a}_i, \breve{b}_i\big)$-process, we still have $\breve{\lambda}_1 = \big(\sqrt{2} - 1\big)^2$ with a similar eigenfunction $\breve{g}$: $\breve{g}_0 = -1$ and

$$\breve{g}_i = \frac{1}{10} 2^{i/2}\Big[-67 + 42\sqrt{2} + \Big(27 - 17\sqrt{2}\Big)\, i\Big], \qquad i \geqslant 1.$$

Now, as an application of Proposition 9.11 (2) with the original (c_i) replacing $c_2 = 4/3$ by $c_2 = 3/2$ only, the resulting λ_0 has a lower bound $\big(\sqrt{2} - 1\big)^2 \approx 0.17$.

(d) To apply Corollary 9.9, we write $c_i = 13(1 + (-1)^i)/12$ and use (9.20) and (9.21):

$$\xi_\varepsilon = \inf_{i \geqslant 2} \frac{1 - \varepsilon + \varepsilon x_i}{z_i + \varepsilon y_i}, \ \ \zeta_\varepsilon = \inf_{\text{odd } i \geqslant 3} \frac{\varepsilon\, \eta}{1 + \varepsilon x_i - \eta(z_i + \varepsilon y_i)}, \ \ \eta \in [0, \xi_\varepsilon], \tag{9.33}$$
$$x_i = \frac{13}{72}\big(2^{2+i} - 6\, i - 3 - (-1)^i\big), \quad y_i = 2^{i-1} - i, \quad z_i = \frac{2^i - 2}{5}.$$

Note that the numerator of ζ_ε given in (9.33) is independent of i but in the denominator, x_i, y_i and z_i all tend to infinity as $i \to \infty$. To avoid the trivial

estimate, one needs to cancel the leading term in i of $\varepsilon x_i - \eta(z_i + \varepsilon y_i)$ in the denominator. This leads to the following solution:

$$\eta = \frac{65\,\varepsilon}{9(2+5\,\varepsilon)}.$$

Inserting this into $\varepsilon x_i - \eta(z_i + \varepsilon y_i)$, it follows that the denominator of ζ_ε in (9.33) becomes

$$1 - \frac{65\left(-2\,i + (-1)^i + 3\right)\varepsilon^2 + 2\left(78\,i + 13(-1)^i - 245\right)\varepsilon - 144}{72(5\,\varepsilon+2)}.$$

Now, in order to remove the leading term in i, the only solution is

$$\varepsilon = 78/65 > 1,$$

which does not belong to the domain of $\varepsilon \in (0,1)$. Therefore, the test function used in Corollary 9.9 does not provide enough freedom to cover this example.

Note that without the killing rate, the process with rates (a_i) and (b_i) is exponentially ergodic and so $\lambda_0(a_i, b_i, 0) = 0$. $\square$

For the following examples, we assume that $a_i = b_i$ for $i \geqslant 2$. Then

$$\mu_1 = 1, \quad \mu_i = b_1 a_i^{-1}, \quad i \geqslant 2; \quad \mu_i b_i = b_1, \quad i \geqslant 1.$$

The quantities ξ_ε and ζ_ε defined in (9.20) and (9.21), respectively, are now determined by

$$x_i = \sum_{2\leqslant j\leqslant i-1} \frac{i-j}{a_j}\tilde{c}_j, \qquad y_i = \sum_{2\leqslant j\leqslant i-1} \frac{i-j}{a_j}, \qquad z_i = \frac{i-1}{b_1}.$$

Example 9.19. *Let $a_1 = 0$, $b_1 = 2\beta(1-\beta)(1-2\beta)^{-1}$ ($\beta \in (0,1/2)$), $a_i = b_i = \beta\, i$ for $i \geqslant 2$, $c_i = (1-\beta)^2(i-1)$ for $i \geqslant 1$. Then $\lambda_0 = 2\beta(1-\beta)$. In the special case that $\beta = 1/4$, we have $\lambda_0 = 3/8$. The upper and lower bounds provided by (9.9) and Corollary 9.9 are* $3/4$ *and approximately* 0.274, *respectively.*

Proof. Let $v_i = \beta(1 + i^{-1})$ for $i \geqslant 1$. Then $R_i(v) \equiv 2\beta(1-\beta)$. By Kummer's test (cf. (9.32)), the corresponding function g defined by (9.31) belongs to $L^2(\mu)$. Hence, the assertion follows from Theorem 9.3. Note that $v_i < 1$ for all i, and g is strictly decreasing even though $c_1 = 0 < \lambda_0$ and $c_i > \lambda_0$ for all $i > (1+\beta)(1-\beta)^{-1}$ (compare with (2.5)).

As an application of (9.9) with $(\ell, m) = (1,1)$ or $(3,4)$, we obtain

$$\lambda_0 \leqslant (1-\beta)\left\{\frac{2\beta}{1-2\beta} \bigwedge \frac{23 - 40\beta + 23\beta^2}{2(8-11\beta)}\right\}.$$

To study the lower bound, for simplicity, we let $\beta = 1/4$. Then $\lambda_0 = 3/8$ and the upper bound in the last formula is $3/4$. Choose $\varepsilon = \left(\sqrt{409} - 5\right)/24$ so that the infimum $\xi_\varepsilon = \left(29 - \sqrt{409}\,\right)/32 \approx 0.274$ is attained simultaneously at $i = 2$ and $i = 3$. Since $\xi_\varepsilon < c_2$, the set $\{i \geqslant 2 : c_i < \xi_\varepsilon\}$ is empty. Therefore, the lower bound provided by Corollary 9.9 is approximately 0.274. $\square$

For the following two examples, without the killing rate, the process is exponentially ergodic and so $\lambda_0(a_i, b_i, 0) = 0$.

Example 9.20. *Let* $a_1 = 0$, $b_1 = 4/5$, $a_i = b_i = i^2$ *for* $i \geqslant 2$, *and*

$$c_i = \frac{8}{9}\left[\frac{8}{3\,i-8} - \frac{2}{3\,i-4} + 5\right], \qquad i \geqslant 1.$$

Then $\lambda_0 = 4$. *The upper and lower bounds provided by (9.9) and Corollary 9.9 are* $14/3$ *and approximately* 2.82, *respectively.*

Proof. The proof is similar as before using

$$v_i = 1 - \frac{1}{3\,i-4}, \qquad i \geqslant 1.$$

Note that c_i has minimum 0 at $i = 2$. The upper bound provided by (9.9) with $(\ell, m) = (2, 2)$ is $14/3$. The lower bound produced by Corollary 9.9 with $\varepsilon = 1$ is $48/17 \approx 2.82$. Since $c_1 > 0$, the parameter $\varepsilon = 1$ is allowed. Then $\xi_\varepsilon = 48/11$ is attained at $i = 3$, and $\zeta_\varepsilon = 48/17$ is attained at $i = 2$ (noting that the set $\{i \geqslant 2 : c_i < \xi_\varepsilon\}$ is a singleton $\{2\}$). $\square$

Example 9.21. *Let* $a_1 = 0$, $b_1 = 3/2$, $c_1 = 15$,

$$a_i = b_i = i\,(i - 4^{-1})(12\,i^2 - 31i + 27), \qquad i \geqslant 2,$$
$$c_i = i^4 - \frac{1}{2}\,i^3 - \frac{301}{16}\,i + \frac{227}{8}, \qquad i \geqslant 2.$$

Then $\lambda_0 = 119/8 = 14.875$. *The upper and lower bounds provided by (9.9) and Corollary 9.9 are approximately* 15.42 *and* 13.18, *respectively.*

Proof. Note that c_i is convex and has its minimum 0 at $i = 2$. For

$$v_i = \frac{3}{4} - \frac{2}{i} + \frac{7}{4\,i-1}, \qquad i \geqslant 1,$$

we have $R_i(v) \equiv 119/8$. Note that $v_1 > 1$ and $v_i < 1$ for all $i \geqslant 2$. The function g defined by (9.31) is not monotone but is bounded. Next, since $\mu_i \sim i^{-4}$, we have $g \in L^2(\mu)$. The assertion now follows from Theorem 9.3.

Clearly, $\inf_{i \geqslant 1} c_i = 11/4$. The upper bound provided by (9.9) with $(\ell, m) = (2, 4)$ is approximately 15.42. To get a lower estimate, we apply Corollary 9.9. Because $\tilde{c}_1 > 0$, we can choose $\varepsilon = 1$. Then $\xi_\varepsilon = 354679/29504$ is attained at $i = 4$. Next, since the set $\{i \geqslant 2 : c_i < \xi_\varepsilon\}$ is a singleton $\{2\}$, we need only to compute ζ_ε at $i = 2$: $\zeta_\varepsilon \approx 10.43$. Thus, the lower bound produced by Corollary 9.9 is approximately 13.18. $\square$

To conclude this section, we return to the uniqueness problem for birth–death processes with killing of the Dirichlet form as discussed at the end of Section 1. Certainly, the problem is meaningful only if $N = \infty$. Recall that for a given Q-matrix, not necessarily conservative (i.e., may have killing), the exit space $\mathscr{U}_\lambda$ is the set of the solutions (u_i) to the following equation:

$$\begin{cases} (\lambda I - Q)u = 0, \\ 0 \leqslant u \leqslant 1, \end{cases} \qquad \lambda > 0.$$

Note that the dimension of $\mathscr{U}_\lambda$ is independent of $\lambda > 0$. By (2.5) replacing λ with $-\lambda$, it follows that the non-trivial exit solution, if it exists, is unique and is strictly increasing.

Theorem 9.22 (Uniqueness of the Dirichlet form). *Let $N = \infty$.*

(1) *The Dirichlet form satisfying the Kolmogorov's equations is unique if $\mathscr{U}_\lambda = \{0\}$. Equivalently,*

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^{n} \mu_k (1 + c_k) = \infty. \tag{9.34}$$

(2) *Let $\sum_{i\in E} \mu_i < \infty$. Then the Dirichlet form is unique iff*

$$\sum_{i\in E} \mu_i c_i < \infty \quad \text{and} \quad \sum_{i\in E} \frac{1}{\mu_i b_i} = \infty.$$

(3) *Let $\sum_{i\in E} \mu_i = \infty$. Then the Dirichlet form is unique if*

$$\text{either } \inf_{i\in \overset{\bullet}{E}} \sum_{j\in E} P_{ij}^{\min}(\lambda) > 0 \quad \text{or} \quad \sum_{i\in E} \mu_i c_i < \infty$$

holds, where $\overset{\bullet}{E} = \{i \in E : c_i > 0\}$.

Here are some comments about the theorem.

(i) Suppose that only a finite number of c_i are non-zero.

Then condition (9.34) is equivalent to (1.2) [Certainly in this item, we are using the modified (1.2) and (1.3) by removing 0 from the state space]:

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^{n} \mu_k \leqslant \sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^{n} \mu_k (1 + c_k) \leqslant C \sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^{n} \mu_k,$$

where $C = \max_{i: c_i > 0} (1 + c_i) < \infty$. Hence, condition (9.34) is stronger than (1.3). In this case, condition (9.34) is even not needed in part (3) where the last two conditions are automatic.

When $\sum_i \mu_i < \infty$, (1.3) is equivalent to (1.2) which coincides with (9.34). When $\sum_i \mu_i = \infty$, both (1.3) and part (3) hold. Therefore, if only a finite number of c_i are non-zero, then we have

$$\text{criterion (1.3)} \iff \text{one of parts (2) and (3) holds.}$$

(ii) When $c_i \neq 0$ for infinite number of i, except condition (9.34), an additional condition on the killing rates (c_i) is required. The condition means that if c_i increases very fast, then there exist some Dirichlet forms that do not satisfy the Kolmogorov equations.

(iii) The second condition in part (3) is the same as the one in part (2). For the first condition in part (3), it is easy to write down some more explicit sufficient conditions. This is due to the following fact. Since for each fixed j, $\{P_{ij}^{\min}(\lambda) : i \in E\}$ is the minimal solution to the equations

$$x_i = \sum_{k\neq i} \frac{q_{ik}}{\lambda + q_i} x_k + \frac{\delta_{ij}}{\lambda + q_i}, \qquad i \in E,$$

by the linear combination theorem, $\big\{\sum_{j\in E} P_{ij}^{\min}(\lambda) : i\in E\big\}$ is the minimal solution to the equations

$$x_i = \sum_{k\neq i} \frac{q_{ik}}{\lambda+q_i} x_k + \frac{1}{\lambda+q_i}, \qquad i\in E.$$

This minimal solution (x_i^*) can be obtained in the following way. Let

$$\begin{aligned} x_i^{(1)} &= \frac{1}{\lambda+q_i}, \qquad i\in E,\\ x_i^{(n+1)} &= \sum_{k\neq i} \frac{q_{ik}}{\lambda+q_i} x_k^{(n)} + \frac{1}{\lambda+q_i}, \qquad i\in E,\ n\geqslant 1. \end{aligned}$$

Then $x_i^{(n)} \uparrow x_i^*$ as $n\to\infty$ for every $i\in E$ (cf. [10; §2.1 and §2.2]). Hence, for each $n\geqslant 1$, $x_i^{(n)}$ is a lower bound of $\sum_{j\in E} P_{ij}^{\min}(\lambda)$. From this discussion, it is clear that the first condition in part (3) is also a restriction on the growing of the killing rates (c_i). This is consistent with the second condition there. It is regretted that we do not know at the moment whether the conditions in part (3) are necessary or not.

Proof of Theorem 9.22. Part (1) follows from [10; Theorem 3.2] and Chen et. al. (2005)[1] with a fictitious state 0. The last cited result is an application of the single birth processes. Noting that if $\sum_{i\in E}\mu_i<\infty$ and $\sum_{i\in E}\mu_i c_i<\infty$, then (9.34) holds iff $\sum_{i\in E}(\mu_i b_i)^{-1}=\infty$, hence, part (2) is a special case of [10; Theorem 6.42]. Next, noting that the unique exit solution is strictly increasing, when $\sum_i \mu_i=\infty$, we have $\mathscr{U}_\lambda\cap L^1(\mu)=\{0\}$. Hence, part (3) is a particular application of [10; Theorem 6.41]. $\square$

10. Notes

10.1 Open problems and basic estimates for diffusions.

Having seen such a long paper, the reader may feel strange if we claim that the story is still incomplete even in the context of birth–death processes. Unfortunately, it is the case.

All of the examples we have done so far show that the following facts hold.

(1) The ratio of the improved upper and lower bounds belongs to $[1,2]$.
(2) The sequence $\{\bar\delta_n\}$ is increasing in n and $\bar\delta_n\geqslant\delta_n'$ for all n.
(3) The sequences $\{\bar\delta_n\}$, $\{\delta_n'\}$, and $\{\delta_n\}$ all converge to λ_0^{-1} as $n\to\infty$.
(4) The relation $(\bar\eta_1,\eta_1)\subset(\kappa,4\kappa)$ discussed in Section 6 holds.

However, there is still no analytic proof for them. The difficulty for the first question is that the maximum/minimum of $\bar\delta_1$ and δ_1 may locate in different places. In the case that (2) would be true, then the story could be simplified since we need the first sequence only. For Questions (2) and (3), the assertions are numerically justified for almost all of the examples in the paper but the results are not included. We have not worked on Question (3) hardly enough since one can go ahead only in a finite number of steps in the symbol computation but

the question is certainly meaningful and in the numerical computation, only in a few steps one achieves the eigenvalue. For the sequences $\{\bar{\eta}_n\}$ and $\{\eta_n\}$, we have similar questions as (1) and (3) about, but the corresponding question (2) is answered by Lemma 6.5.

There is a parallel story for the one-dimensional diffusions. In many cases, one can easily guess what the result should be, even though there may exist a new difficulty in its proof. For instance, as a combination of the proofs of Theorem 8.2 and [12; Corollary 7.6], one may prove the following result.

Theorem 10.1. *Consider the minimal diffusion on $(-M, N)\,(M, N \leqslant \infty)$ with operator*

$$L = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x} \qquad \left(a(x) > 0,\ \frac{b(x)}{a(x)} \text{ is locally integrable}\right),$$

and Dirichlet boundaries at $-M$ if $M < \infty$, and at N if $N < \infty$. Let $C(x) = \int_\theta^x b/a$ for some fixed reference point $\theta \in (-M, N)$ and assume additionally that e^C/a is also locally integrable. Denote by $A_{\mathbb{B}}$ the optimal constant in Poincaré-type inequality (8.1) with Dirichlet form

$$D(f) = \int_{-M}^{N} f'^2 e^C, \qquad f \in \mathscr{C}_0^\infty(-M, N),$$

Then $A_{\mathbb{B}}$ satisfies $B_{\mathbb{B}} \leqslant A_{\mathbb{B}} \leqslant 4B_{\mathbb{B}}$, where

$$B_{\mathbb{B}}^{-1} = \inf_{-M<x<y<N} \left[\left(\int_{-M}^{x} e^{-C}\right)^{-1} + \left(\int_{y}^{N} e^{-C}\right)^{-1}\right] \|\mathbb{1}_{(x,y)}\|_{\mathbb{B}}^{-1}. \tag{10.1}$$

By the way, we prove a dual result of Theorem 10.1 for ergodic diffusions. As discussed in the proof of Theorem 7.5, the exponentially ergodic rate often coincides with the first non-trivial eigenvalue λ_1 defined below. Consider a diffusion process with operator L as in Theorem 10.1, with state space $(-M, N)\,(M, N \leqslant \infty)$ and reflecting boundaries at $-M$ if $M < \infty$, and at N if $N < \infty$. For convenience, we define two measures as follows:

$$\text{Scale measure: } \nu(\mathrm{d}x) = e^{-C(x)}\mathrm{d}x, \qquad C(x) := \int_\theta^x \frac{b}{a},$$

where $\theta \in (-M, N)$ is a fixed reference point.

$$\text{Speed measure: } \mu(\mathrm{d}x) = \frac{e^{C(x)}}{a(x)}\,\mathrm{d}x.$$

With these measures, the operator L takes a compact form:

$$L = \frac{\mathrm{d}}{\mathrm{d}\mu}\frac{\mathrm{d}}{\mathrm{d}\nu}. \tag{10.2}$$

Next, suppose that $\mu(-M,N)<\infty$, and denote by π the normalized probability measure of μ. Set

$$\mathscr{A}=\{f: f \text{ is absolutely continuous in } (-M,N)\},$$

and define

$$\lambda_1=\inf\{D(f): f\in L^2(\mu)\cap\mathscr{A},\ \pi(f)=0,\ \|f\|=1\},$$

where

$$D(f)=\int_{-M}^N a f'^2 \mathrm{d}\mu, \qquad f\in\mathscr{A}.$$

Clearly, in the definition of λ_1, only those f in the set $\{f\in L^2(\mu)\cap\mathscr{A}:\ D(f)<\infty\}$ are useful. In other words, we are here using the maximal Dirichlet form, as in Section 6.

Theorem 10.2. *Let $a>0$, a and b be continuous on $[-M,N]$ (or $(-M,N]$ if $M=\infty$, for instance). Assume that $\mu(-M,N)<\infty$. Then for λ_1, we have the basic estimate: $\kappa^{-1}/4\leqslant\lambda_1\leqslant\kappa^{-1}$, where*

$$\kappa^{-1}=\inf_{-M<x<y<N}\bigg[\bigg(\int_{-M}^x \mathrm{d}\mu\bigg)^{-1}+\bigg(\int_y^N \mathrm{d}\mu\bigg)^{-1}\bigg]\bigg(\int_x^y \mathrm{d}\nu\bigg)^{-1}. \tag{10.3}$$

Proof. (a) First we show that for the basic estimate, it suffices to consider the finite M and N with smooth a and b. Since a and b are continuous, if $M=N=\infty$ for instance, we may choose $M_p,\ N_p\uparrow\infty$ as $p\to\infty$ such that $\theta\in(-M_p,N_p)$ for all p. Then, by Chen and Wang (1997, Lemma 5.1), we have $\lambda_1^{(M_p,N_p)}\downarrow\lambda_1$ as $p\to\infty$ (This is parallel to the localizing procedure used in Section 6). At the same time, the isoperimetric constants $\kappa^{(M_p)^{-1}}\downarrow\kappa^{-1}$ as $p\to\infty$ (cf. proof of Corollary 7.9). Hence, in what follows, we may assume that $M,N<\infty$. Next, by using the continuity of a and b again, and using a standard smoothing procedure, we can choose smooth a_p and b_p such that $a_p\to a$ and $b_p\to b$ (as $p\to\infty$) uniformly on finite intervals, and furthermore, we can assume that $a_p>0$ on each fixed closed finite interval. Clearly, the corresponding κ_p converges to κ as $p\to\infty$. Therefore, without loss of generality, we assume, unless otherwise stated, that not only $M,N<\infty$ but also a and b are smooth with $a>0$ on $[-M,N]$.

(b) Recall the following differential form of variational formula for λ_1:

$$\lambda_1=\sup_{f\in\mathscr{F}}\inf_{x\in(-M,N)}\bigg[-b'-\frac{af''+(a'+b)f'}{f}\bigg](x), \tag{10.4}$$

where

$$\mathscr{F}=\big\{f\in\mathscr{C}^1[-M,N]\cap\mathscr{C}^2(-M,N): f(-M)=f(N)=0,\ f|_{(-M,N)}>0\big\}.$$

This is an analog of the variational formula for the lower estimate in Theorem 6.1 (1). In the original study by Chen and Wang (1997, (2.3)), the state

space is the half-line, not finite, but this is not essential. It works also for finite state spaces. Besides, it was stated as "$\geqslant$" in (10.4) only. For "=", one simply chooses $f = g'$, where g is the eigenfunction of λ_1. This gives us the boundary condition: $f(-M) = f(N) = 0$ since $g'(-M) = g'(N) = 0$ by assumption. Here, one requires that $g \in \mathscr{C}^3(-M, N)$ which is satisfied since we are now in a finite interval having smooth a and b. Alternatively, instead of the original coupling proof, one may use the analytic one which leads to (6.4) for birth–death processes.

(c) We are now going to handle with a more general situation: $M, N \leqslant \infty$, $a, b \in \mathscr{C}^1(-M, N)$ and $a > 0$ on $(-M, N)$. Let us define a dual operator $\widehat{L}$ of L. In view of the Karlin and McGregor's construction, the dual of a birth–death process is simply an exchange of the scale and speed measures $\hat{\mu} = \nu$ and $\hat{\nu} = \mu$ up to a constant (cf. (5.3)). Thus, in view of (10.2), the dual operator $\widehat{L}$, as was introduced by Cox and Rösler (1983), should be given by

$$\widehat{L} = \frac{\mathrm{d}}{\mathrm{d}\hat{\mu}}\frac{\mathrm{d}}{\mathrm{d}\hat{\nu}}. \tag{10.5}$$

Again, the speed and scale measures $\hat{\mu}$ and $\hat{\nu}$ of $\widehat{L}$ should be expressed as

$$\mathrm{d}\hat{\mu} = \frac{e^{\widehat{C}}}{\hat{a}}\mathrm{d}x, \qquad \mathrm{d}\hat{\nu} = e^{-\widehat{C}}\mathrm{d}x$$

in terms of the coefficients $\hat{a}$ and $\hat{b}$ of $\widehat{L}$ to be determined now. Because $\hat{\mu} = \nu$ and $\hat{\nu} = \mu$, we have

$$\frac{\mathrm{d}\hat{\mu}}{\mathrm{d}x}\frac{\mathrm{d}\hat{\nu}}{\mathrm{d}x} = \frac{\mathrm{d}\nu}{\mathrm{d}x}\frac{\mathrm{d}\mu}{\mathrm{d}x}.$$

It follows that $\hat{a} = a$. Then using the equation $\hat{\mu} = \nu$, we get

$$\widehat{C} = -C + \log \hat{a} = -C + \log a.$$

Thus, from

$$\frac{\hat{b}}{\hat{a}} = \widehat{C}' = -\frac{b}{a} + \frac{a'}{a},$$

we get $\hat{b} = a' - b$. Therefore, the dual operator $\widehat{L}$ has the following expression:

$$\widehat{L} = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \left(\frac{\mathrm{d}}{\mathrm{d}x}a(x) - b(x)\right)\frac{\mathrm{d}}{\mathrm{d}x}. \tag{10.6}$$

For the dual process, the Dirichlet boundary is endowed at $-M$ and N (cf. proof (e) below). Clearly, the dual operator $\widehat{L}$ is symmetric on $L^2\big(e^{-C}\mathrm{d}x\big)$. We remark that the assumption on a and b can be weakened in this paragraph.

(d) Define a Schrödinger operator as follows:

$$\begin{aligned} L_S &= a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \big(a'(x) + b(x)\big)\frac{\mathrm{d}}{\mathrm{d}x} + b'(x) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(a(x)\frac{\mathrm{d}}{\mathrm{d}x}\right) + b(x)\frac{\mathrm{d}}{\mathrm{d}x} + b'(x), \end{aligned} \tag{10.7}$$

with Dirichlet boundaries at $-M$ and N provided they are finite. Clearly, L_S is symmetric on $L^2\big(e^C\mathrm{d}x\big)$. Denote by λ_S the principal eigenvalue of L_S:

$$\lambda_S = \bigg\{ -(f, L_S f)_{L^2(e^C\mathrm{d}x)} : f \in \mathscr{C}_0^\infty(-M,N),\ \int_{-M}^N f^2 e^C = 1\bigg\}.$$

In the setup of (b), formula (10.4) becomes

$$\lambda_1 = \sup_{f\in\mathscr{F}} \inf_{x\in(-M,N)} \frac{-L_S f}{f}(x).$$

This leads to the study on λ_S.

(e) An elementary computation shows that

$$e^C L_S e^{-C} = \widehat{L}. \tag{10.8}$$

Note that

$$\int_{-M}^N e^C f L_S g = \int_{-M}^N e^{-C}(e^C f)\big(e^C L_S e^{-C}\big)(e^C g) = \int_{-M}^N e^{-C} \hat{f}\widehat{L}\hat{g},$$

where the mapping $f \to \hat{f} := e^C f$ is an isometry from $L^2\big(e^C\mathrm{d}x\big)$ to $L^2\big(e^{-C}\mathrm{d}x\big)$, and that $\hat{f} \in \mathscr{C}_0^2(-M,N)$ iff $f \in \mathscr{C}_0^2(-M,N)$. Since $\mathscr{C}_0^2(-M,N)$ is also a common core of L_S and $\widehat{L}$ by the assumption on the coefficients a and b, it follows that the operators L_S and $\widehat{L}$ with the same core $\mathscr{C}_0^\infty(-M,N)$ are isospectral. In particular, we have $\lambda_S = \hat{\lambda}_0$. When $M, N < \infty$, this means that $\widehat{L}$ has Dirichlet boundaries at $-M$ and N since so does L_S. Now, the basic estimates for λ_S can be obtained in terms of the ones for $\hat{\lambda}_0$, as will be shown in part (f) below.

To go back to λ_1, noting that by (10.8) again, we also have

$$\frac{-L_S f}{f} = \frac{-(e^C L_S e^{-C})(e^C f)}{e^C f} = \frac{-\widehat{L}\hat{f}}{\hat{f}}.$$

By (a), we can assume that $M, N < \infty$ and $a > 0$ on $[-M,N]$. From Shiozawa and Takeda (2005) and X. Zhang (2007), it is known that

$$\hat{\lambda}_0 \geqslant \sup_{f\in\mathscr{F}} \inf_{x\in(-M,N)} \frac{-\widehat{L}\hat{f}}{\hat{f}}(x)$$

(i.e., Barta's inequality). To see that the equality sign holds, simply choose f to be the eigenfunction $\hat{g}$ of $\hat{\lambda}_0$. The fact that $\hat{g} \in \mathscr{F}$ is guaranteed by the assumptions that $M, N < \infty$, $\hat{a}$ and $\hat{b}$ are continuous, and $\hat{a} > 0$ on $[-M,N]$. This is a standard (regular) Sturm–Liouville eigenvalue problem. The property $\hat{g}|_{(-M,N)} > 0$ is due to the fact that $\hat{\lambda}_0$ is the minimal eigenvalue. We have thus returned to λ_1 from $\hat{\lambda}_0$ through λ_S.

(f) For the dual operator $\widehat{L}$ defined in part (c), applying Theorem 10.1 to $\mathbb{B} = L^1(\hat{\mu})$, we obtain $\hat{\kappa}^{-1}/4 \leqslant \hat{\lambda}_0 \leqslant \hat{\kappa}^{-1}$, where

$$\hat{\kappa}^{-1} = \inf_{-M<x<y<N} \bigg[\bigg(\int_{-M}^{x} \mathrm{d}\hat{\nu}\bigg)^{-1} + \bigg(\int_{y}^{N} \mathrm{d}\hat{\nu}\bigg)^{-1}\bigg]\bigg(\int_{x}^{y} \mathrm{d}\hat{\mu}\bigg)^{-1}.$$

Now, the theorem follows by the dual transform $\hat{\mu} = \nu$ and $\hat{\nu} = \mu$.

Finally, the proof of Theorem 10.2 can be summarized as follows:

λ_1 for general M, N and continuous a, b
$\to \lambda_1$ for finite M, N and smooth a, b
(by approximating and smoothing procedure)
$\to \lambda_S$ (by coupling method leading to the Schrödinger operator)
$\to \hat{\lambda}_0$ (by isometry in terms of the dual operator)
$\to$ basic estimate of $\hat{\lambda}_0$ (by capacitary method: Theorem 10.1)
$\to$ basic estimate of λ_1 (by duality). $\square$

Actually, we have also proved the following result (cf. parts (c)–(f) in the last proof) which is parallel to Proposition 9.11.

Proposition 10.3. *Let $M, N \leqslant \infty$, $a, b \in \mathscr{C}^1(-M, N)$ and $a > 0$ on $(-M, N)$. Then for the Schrödinger operator L_S on $L^2\big(e^C \mathrm{d}x\big)$ having the form (10.7) with Dirichlet boundaries at $-M$ if $M < \infty$, and at N if $N < \infty$, we have $\lambda_S = \hat{\lambda}_0$, and furthermore, $\kappa^{-1}/4 \leqslant \lambda_S \leqslant \kappa^{-1}$, where κ is defined by (10.3).*

The following simplified estimate of $\kappa^{(10.3)}$ is helpful in practice. Recall that by assumption, $\mu(-M, N) < \infty$. Let $m(\mu)$ be the median of μ (i.e., $\mu(-M, m(\mu)) = \mu(m(\mu), N)$). Given $x \in (-M, m(\mu))$, let $y = y(x)$ be the unique solution to the equation: $\mu(y, N) = \mu(-M, x)$. The A-G inequality $\alpha + \beta \geqslant 2\sqrt{\alpha\beta}$ suggests the use of $y(x)$, which then leads to a simpler bound:

$$\kappa^{(10.3)} \geqslant 2^{-1} \sup_{x \in (-M,\, m(\mu))} \mu(-M, x)\, \nu(x, y(x)).$$

We remark that the equality sign here holds in some cases, but the inequality sign can happen in general. Anyhow, this provides us a guidance in seeking for the infimum in (10.3). Certainly, the similar discussion is meaningful for $\kappa^{(10.1)}$.

Having Theorems 10.1 and 10.2 at hand, the basic estimates in the other cases (λ^{ND} and λ^{DN}) mentioned in Section 1 should be clear.

The study on the one-dimensional case provides a comparison tool for the study on the higher dimensional situation, as we did a lot before. Hence, there is no doubt for the development in the higher dimensional context.

10.2 *h*-Transform.

In an earlier draft of this paper (roughly speaking, up to Theorem 7.1 plus a part of Theorem 9.3), the author mentioned an open question: how to handle

the case that (1.3) fails? Then two answers have appeared. The first one is the use of so-called h-transform by Wang (2008a) where the transient case studied in Section 7 is transferred into the one studied in Section 4. Next, with the help of the duality given in Theorem 7.1, the ergodic case studied in Section 6 can be also transferred into the one studied in Section 4. In this way, with a use of Theorem 4.2, Wang obtains a criterion for λ_1 (Section 6) with a factor 4. To have a taste of this technique, let us quote a particular result here.

Theorem 10.4 (Wang (2008a, Theorem 1.2)). *Set $h_i = \sum_{j=i}^N \mu_j$. Then we have $\delta^{-1}/4 \leqslant \lambda_1 \leqslant \delta^{-1}$, where*

$$\delta = \sup_{1 \leqslant i < N+1} \left(\frac{1}{h_i} - \frac{1}{h_0}\right) \sum_{j=i}^N \frac{1}{\mu_i a_i} h_i^2. \tag{10.9}$$

Comparing this result with Corollary 6.6, the factor 4 is in common but the isoperimetric constants are quite different. The advantage here is that only one variable is required in the supremum, but in Corollary 6.6 two variables are needed. The price one has to pay to (10.9) is involving a new quantity h. The natural extension of Corollary 6.6 to the whole line (Corollary 7.9) exhibits an interesting symmetry of the left and the right half-lines. Such an extension of Theorem 10.4 with the same factor 4 is unclear to the author. Along the same line and using [9], Wang then extends the results to Poincaré-type inequalities as well as functional inequalities, see Wang (2008b, c). Clearly, Wang's papers show that the h-transform is a powerful tool and may be useful in other cases.

While the author's solution to the above open question is the use of the maximal process as included into this version of the paper. As shown in the paper, Corollary 6.6 comes from the author's previous general result without using the h-transform. An interesting question in mind is to use the variational formulas in Section 6 to derive Corollary 6.6 directly. Besides, a direct generalization of Sections 2, 3, and 7 to the Poincaré-type inequalities is still meaningful in practice since the formulas are quite different (in view of Theorem 10.4) and some of them may be more practicable.

10.3 Remark on some known results.

As mentioned in Section 5, duality (5.1) goes back to Karlin and McGregor (1957b). The author learned this technique mainly from van Doorn (1981; 1985) based on which the proof of the basic result $\lambda_1 = \alpha^*$ was done, cf. [2]. It is now known that such a result holds in a very general setup as indicated in the proof of Theorem 7.4.

We now discuss the situation that (1.3) holds. Then there are three cases.

(1) $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i b_i)^{-1} < \infty$.
(2) $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i b_i)^{-1} = \infty$.
(3) $\sum_i \mu_i = \sum_i (\mu_i b_i)^{-1} = \infty$.

First, let $b_0 > 0$. In cases (1) or (3), by Theorem 2.4 (1) and Proposition 2.7 (1),

$\lambda_0^{(2.2)}$ is equal to

$$\sup_{v\in\mathscr{V}}\inf_{i\geqslant 0}\,[a_{i+1}+b_i-a_i/v_{i-1}-b_{i+1}v_i]. \tag{10.10}$$
$$\mathscr{V}:=\{v: v_{-1}\text{ is free, } v_i>0 \text{ for all } i\geqslant 0\}.$$

In case (2), by Theorem 6.1 (1), λ_1 can be expressed by (10.10). Thus, in view of Proposition 1.2 and [2; Theorem 5.3], the convergence rate α^* can be also expressed by (10.10).

Next, let $b_0=0$. Then in case (2), by Corollary 5.2, Proposition 2.7 (1), and using (5.8) in an inverse way, it follows that $\lambda_0^{(4.2)}$ is equal to

$$\sup_{v\in\mathscr{V}}\inf_{i\geqslant 1}\left[a_i\left(1-\frac{1}{v_{i-1}}\right)+b_i(1-v_i)\right], \tag{10.11}$$
$$\mathscr{V}:=\{v: v_0=\infty,\ v_i>0 \text{ for all } i\geqslant 1\}.$$

In case (1), $\lambda_0^{(4.2)}$ is equal to $\lambda_0^{(7.1)}$. By Theorem 7.1 (1), in terms of Theorem 6.1 (1) and using (5.8) in an inverse way, we obtain the same expression (10.11) for $\lambda_0^{(7.1)}$. Finally, in the degenerated case (3), we indeed have $\lambda_0^{(4.2)}=\lambda_0^{(7.1)}=0$ which can be expressed as (10.11) by Theorem 7.1 (2). Hence, by Proposition 1.2, the convergence rate α^* can also be expressed by (10.11). We have thus obtained the following result.

Theorem 10.5 (van Doorn (2002))**.** *Let* (1.3) *hold. Then the exponential convergence rate α^* is given by* (10.10) *or* (10.11)*, respectively, according to $b_0>0$ or $b_0=0$.*

With a slightly different expression, this result was given in van Doorn (2002) by the analysis on the extreme zeros of orthogonal polynomials in Karlin and McGregor's representation, and was actually implied in van Doorn's earlier papers (1985; 1987) as mentioned in the paper just cited or in [3]. In the last paper, this result was rediscovered in the study on λ_1, using the coupling methods. The lower estimate was also obtained by Zeifman (1991) using a different method in the case that the rates of the processes are bounded, with a missing of the equality.

A progress made in the paper is removing Condition (1.3) and even (1.2). In particular, the situation having finite state spaces is included. This is meaningful not only theoretically but also in practice since the infinite situation can be approximated by the finite ones. Besides, when $b_0>0$ and $\sum_i\mu_i<\infty$, the duality given by (5.9) is essentially different from (5.8) (cf. Remark 2.8). From the other point of view, the dual of this case goes to $\lambda_0^{(7.1)}$ rather than $\lambda_0^{(4.2)}$. However, we then have to use the maximal process in Section 6, as we did in Theorem 7.1, rather than the minimal one used in Sections 2 and 3, except using (1.3) (which is equivalent to (1.2) if $\sum_i\mu_i<\infty$). From analytical point of view, the use of the maximal process is natural since one looks for the inequality to be held for the largest class of functions, as illustrated by the weighted Hardy inequality in Section 4.1.

In van Doorn (2002), some variational formulas of difference form for the upper bound of α^* are also presented but we do not use them here. As far as we know, the criterion for $\alpha^* > 0$ (Theorem 1.5) has been open for quite a long time; it was answered in the ergodic case only till [6] in terms of the study on the first non-trivial eigenvalue λ_1. For which, the criterion was obtained independently by Miclo (1999) based on the weighted Hardy's inequality. Criterion 3.1 follows from the variational formulas of single summation form (part (2) of Theorem 2.4), but it is not obvious at all to deduce the criterion from (10.10) (or dually from (10.11)) directly. More clearly, the variational formula of the difference form for the lower bound given in (9.3) which is closely related to (10.11) was known for some years and works in a more general setup, but an explicit criterion for the killing case is still open (Open Problem 9.13). Anyhow, having the duality (Corollary 5.2 and Theorem 7.1) at hand, Theorem 1.3 is essentially known from [6], except the basic estimates in the ergodic case as well as in the setting of Section 7 is presented here for the first time. The technique adopted in this paper depends heavily on the spectral theory, potential theory, and harmonic analysis. In the transient continuous context, Criterion 3.1 was obtained by Maz'ja (1985, §1.3), as a straightforward consequence of Muckenhoupt (1972). The discrete version was proved by Mao (2002, Proposition A.2). In these quoted papers, the problem in a more general (L^q, L^p)-setup was done.

In the continuous context, the Hardy-type or Sobolev-type inequalities (cf. Theorem 10.1) were studied first by P. Gurka and then by Opic and Kufner (1990, Theorem 8.3). Instead of $\mathscr{D}^{\min}(D)$, they considered the following class of functions: the absolutely continuous functions vanishing at $-M$ and N. This seems not essential in view of $\lambda_0^{(2.2)} = \lambda_0^{(2.18)}$. With a different but equivalent isoperimetric constant (i.e., replacing the sum in (10.1) by maximum "$\vee$"), they obtained upper and lower bounds with ratio $2\,\omega^5 \approx 22$, where $\omega = \big(\sqrt{5}+1\big)/2$ is the gold section number. By the way, we mention that the use of weight functions w and v in (8.6) in the cited book is formally more general than our setup. One can first assume that w and v are positive, otherwise replace them by $w + 1/n$ and $v + 1/n$, respectively, and then pass to the limit as $n \to \infty$. Next, it is easy to rewrite w and v as e^C/a and e^C for some functions C and $a > 0$. Note that only C and a (without using b) are needed to deduce the basic estimates in our proof. Again, in the continuous context, the splitting technique was also used in Theorem 8.8 of the book just quoted where some basic estimates were derived in terms of an isoperimetric constant, up to a factor 8. Their isoperimetric constant is parallel to the right-hand side of (7.13) replacing $\lambda_0^{\theta\pm}$ by the corresponding $\delta^{(3.1)\pm}$ depending on θ (certainly, without using the parameter γ). Our Example 8.9 is an analog of Examples 6.13 and 8.16 in the quoted book. In contrast with our probabilistic–analytic proof here, their proof is direct, analytic, and works in a more general (L^q, L^p)-setup. We have not seen the discrete analog of their results in the literature. In the (L^p, L^p)-sense ($p \geqslant 1$), the variational formulas in the continuous context were obtained in Jin (2006) but it remains open for the more general (L^q, L^p)-setup. Even though it is a typical Sturm-Liouville eigenvalue problem having richer literature, we are unable to find an analog of Theorem 10.2.

Finally, in computing the examples in the paper, the author uses the software Mathematica. All the examples were checked by Ling-Di Wang and Chi Zhang using MatLab. Most of the author's papers cited here can be found in [8].

Acknowledgements The first version of the paper dated May 25, 2007. In June, 2008, the author was invited by Professors Tiee-Jian Wu and Yuan-Shih Chow to visit the Center for Theoretical Sciences at Cheng Kung University and Taiwan University. The author was also invited by Professors Mong-Na Lo Haung and Mei-Hui Guo (Sun Yat-sen University), by Professor Wen-Jang Huang (University of Kaohsiung), by Professor Chen-Hai Tsao (Dong Hwa University for the 17^{th} Conference on Statistics in the Southern area and the Annual Meeting of Chinese Institute of Probability and Statistics), by Professors Chii-Ruey Hwang, Tzuu-Shuh Chiang, Yunshyong Chow and Shuenn-Jyi Sheu (Academia Sinica) for visiting their institutes. The scientific communications, the very warm hospitality of the inviters and the financial support from the Center for Theoretical Sciences, and the inviters' universities are deeply acknowledged. The author has also benefited from several discussions with Professor Shuenn-Jyi Sheu in Beijing. A part of the results in the paper were reported during the visit.

Partial results of the paper have also been reported in the following conferences: International Conference on Stochastic Analysis and Related Fields (April 2008, Wuhan), Workshop on Probability (November 2008, Guangzhou), Chinese-French Meeting in Probability and Analysis (September 2008, Marseille), Workshop on IPS in honor of Professor Tom Liggett's 65th Birthday (June 2009, Peking University), IMS-China International Conference on Statistics and Probability (July 2009, Weihai). The author acknowledges the organizers Zhi-Ming Ma and Michael Röckner, Jia-Gang Ren, Dominique Bakry and Yves LeJan, Da-Yue Chen, Jia-An Yan, respectively, for their invitation and financial support.

Because of an honorary professorship from Swansea University, the author was invited to a workshop and visited University of Strathclyde in UK in October, 2009. The recent progress was reported at these two universities. The author appreciates the invitation, the very warm hospitality, and the financial support by Professors N. Jacob, X.R. Mao, F.Y. Wang, J.L. Wu, C.G. Yuan, and their institutes. Special thanks are given to the speakers of the celebration workshop for their kindness: T.J. Lyons, M. Davis, F.Y. Wang, T.S. Zhang, X.R. Mao, and D. Elworthy.

Two drafts of the paper have been reported in a series of seminars organized by Professors Yong-Hua Mao and Yu-Hui Zhang. Especially, the drafts are checked line by line by Zhang's graduate students Ling-Di Wang and Chi Zhang. The corrections and improvement obtained from them improved a lot the quality of the paper. Their serious work is acknowledged. Since May 2007, the author has reported the results from time to time in his seminars and thanks are given to the participants for their discussions and comments.

References

1. Chen, A., Pollett, P., Zhang, H. and Cairns, B. (2005), *Uniqueness criteria for continuous-time Markov chains with general transition structure*, Adv. Appl. Prob. 37:4, 1056–1074.

2. Chen, M.F. (1991), *Exponential L^2-convergence and L^2-spectral gap for Markov processes*, Acta Math. Sin., New Series 7:1, 19-37.
3. Chen, M.F. (1996), *Estimation of spectral gap for Markov chains*, Acta Math. Sin., New Series 12:4, 337-360.
4. Chen, M.F. (1999), *Analytic proof of dual variational formula for the first eigenvalue in dimension one*, Sci. in China (A) 42:8, 805–815.
5. Chen, M.F. (2000a), *The principal eigenvalue for jump processes*, Acta Math. Sin. Eng. Ser. 16:3, 361–368.
6. Chen, M.F. (2000b), *Explicit bounds of the first eigenvalue*, Sci. Chin. Ser. A 43(10), 1051–1059.
7. Chen, M.F. (2001), *Variational formulas and approximation theorems for the first eigenvalue in dimension one*, Sci. Chin. Ser. A 44(4), 409–418.
8. Chen, M.F. (2001–), *Ergodic Convergence Rates of Markov Processes — Eigenvalues, Inequalities and Ergodic Theory* [Collection of papers, 1993—] http://math.bnu.edu.cn/~chenmf/main_eng.htm.
9. Chen, M.F. (2003), *Variational formulas of Poincaré-type inequalities for birth-death processes*, Acta Math. Sin. Eng. Ser. 19:4, 625-644.
10. Chen, M.F. (2004), *From Markov Chains to Non-equilibrium Particle Systems*, World Scientific. 2nd ed. (1st ed., 1992).
11. Chen, M.F. (2005a), *Capacitary criteria for Poincaré-type inequalities*, Potential Theory 23:4, 303–322.
12. Chen, M.F. (2005b), *Eigenvalues, Inequalities, and Ergodic Theory*, Springer, London.
13. Chen, M.F. Wang, F.Y. (1997), *Estimation of spectral gap for elliptic operators*, Trans. Amer. Math. Soc. 349:3, 1239-1267.
14. Chen, M.F. and Wang, F.Y. (1998), *Cheeger's inequalities for general symmetric forms and existence criteria for spectral gap*, Abstract. Chin. Sci. Bulletin 43:18, 1516–1519. Ann. Prob. 2000, 28:1, 235–257.
15. Chen, M.F., Zhang, Y.H. and Zhao, X.L. (2003), *Dual variational formulas for the first Dirichlet eigenvalue on half-line*, Sci. China 46:6, 847–861.
16. Cox, J.T. and Rösler, U. (1983), *A duality relation for entrance and exit laws for Markov processes*, Stoch. Proc. Appl. 16, 141–156.
17. Dobrushin, R.L. (1952), *On conditions of regularity of stationary Markov processes with a denumerable number of possible states* (in Russian), Uspehi Matem. Nauk (N.S.) 7:6, 185–191.
18. Fukushima, M. and Uemura, T. (2003), *Capacitary bounds of measures and ultracontractivity of time changed processes*, J. Math. Pure et Appliquees 82:5, 553-572.
19. Hou, Z.T., Liu, Z.M., Zhang, H.J., Li, J.P., Zhou, J.Z., and Yuan, C.G. (2000), *Birth–death Processes* (in Chinese), Hunan Sci. Press, Hunan.
20. Hou, Z.T., Zhou, J.Z., Zhang, H.J., Liu, Z.M., Xiao, G.N., Chen, A.Y., and Fei, Z.L. (1994), *The Q-matrix Problem for Markov Chains* (in Chinese), Hunan Sci. Press, Hunan.
21. Jin, H.Y. (2006), *Estimating the constant in L^p-Poincaré inequality* (in Chinese), Master Thesis, Beijing Normal Univ.
22. Karlin, S. and McGregor, J.L. (1957a), *The differential equations of birth-and-death processes, and the Stieltjes moment problem*, Trans. Amer. Math. Soc. **85**, 589-646.
23. Karlin, S. and McGregor, J. (1957b), *The classification of birth and death processes*, Trans. Amer. Math. Soc. 86:2, 366-400.
24. Kijima, M. (1997), *Markov Processes for Stochastic modeling*, Chapman & Hall, London.
25. Mao, Y.H. (2002), *Nash inequalities for Markov processes in dimension one*, Acta. Math. Sin. Eng. Ser. 18:1: 147–156.
26. Mao, Y.H. and Xia, L.Y. (2009), *Spectral gap for jump processes by decomposition method*, Front. Math. China 4:2, 335–347.
27. Maz'ya, V.G. (1985), *Sobolev Spaces*, Springer.
28. Miclo, L. (1999), *An example of application of discrete Hardy's inequalities*, Markov Processes Relat. Fields 5, 319–330.
29. Muckenhoupt, B. (1972), *Hardy's inequality with weights*, Studia Math. XLIV: 31–38.
30. Opic, B. and Kufner, A. (1990), *Hardy-type Inequalities*, Longman, New York.

31. Shao, J.H. and Mao, Y.H. (2007), *Estimation of the Dirichlet eigenvalue of birth–death process on trees*, Acta Math. Sin. Chinese Ser. 50:3, 507–516.
32. Shiozawa, Y. and Takeda, M. (2005), *Variational formula for Dirichlet forms and estimates of principal eigenvalues for symmetric α-stable processes*, Potential Analysis **23**, 135–151.
33. Sirl, D., Zhang, H. and Pollett, P. (2007), *Computable bounds for the decay parameter of a birth-death process*, J. Appl. Prob. 44:2, 476-491.
34. van Doorn, E.A. (1981), *Stochastic Monotonicity and Queuing Applications of Birth–Death Processes*, Lecture Notes in Statistics vol. 4, Springer-Verlag.
35. van Doorn, E.A. (1985), *Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process*, Adv. Appl. Prob. **17**, 514–530.
36. van Doorn, E.A. (1987), *Representations and bounds for zeros of orthogonal polynomials and eigenvalues of sign-symmetric tri-diagonal matrices*, J. Approx. Th. 51, 254–266.
37. van Doorn, E.A. (2002), *Representations for the rate of convergence of birth–death processes*, Theory Probab. Math. Statist. **65**, 37–43.
38. Wang, J. (2008a), *First Dirichlet eigenvalue of transient birth–death processes*, preprint.
39. Wang, J. (2008b), *Poincaré-type inequalities for transient birth–death processes*, preprint.
40. Wang, J. (2008c), *Functional inequalities for transient birth–death processes and their applications*, preprint.
41. Wang, Z.K. (1964), *The ergodicity and zero-one law for birth and death processes* (in Chinese), Acta Sci. Nankai Univ. 5:5, 93–102.
42. Wang, Z.K. and Yang, X.Q. (1992), *Birth and Death Processes and Markov Chains*, Springer, Berlin and Sci. Press, Beijing.
43. Zeifman, A.I. (1991), *Some estimates of the rate of convergence for birth and death processes*, J. Appl. Prob. **28**, 268–277.
44. Zhang, X. (2007), *The estimation for the first eigenvalue of Schrödinger operators and a class of geometric inequalities* (in Chinese), Ph.D Thesis, Beijing Normal University.

School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems (Beijing Normal University), Ministry of Education, Beijing 100875, The People's Republic of China.
E-mail: mfchen@bnu.edu.cn
Home page: http://www.bnu.edu.cn/~chenmf/main_eng.htm